

Propagation of chaos for the Vlasov-Poisson-Fokker-Planck system in 1D

Maxime Hauray and Samir Salem*

Abstract

We consider a particle system in 1D, interacting via repulsive or attractive Coulomb forces. We prove the trajectorial propagation of molecular chaos towards a nonlinear SDE associated to the Vlasov-Poisson-Fokker-Planck equation. We obtain a quantitative estimate of convergence in expectation, with an optimal convergence rate of order $N^{-1/2}$. We also prove some exponential concentration inequalities of the associated empirical measures. A key argument is a weak-strong stability estimate on the (nonlinear) VPFP equation, that we are able to adapt for the particle system in some sense.

Keywords: Particles system, 1D Vlasov-Poisson equation, Propagation of molecular chaos, Monge-Kantorovich-Wasserstein distances, exponential concentration inequalities.

AMS Subject Classification: 35Q83 Vlasov-like equations, 60K35 Interacting random processes, 60H10 Stochastic ordinary differential equations, 82C21 Dynamic continuum models.

1 Introduction

We consider here a one dimensional system of N particles, with position $X_i^N \in \mathbb{R}$ and velocity $V_i^N \in \mathbb{R}$, interacting via the Poisson interaction, and submitted to independent Brownian noises and friction. The associated system of Stochastic Differential Equation (SDE) is the following:

$$dX_{i,t}^N = V_{i,t}^N dt, \quad dV_{i,t}^N = \left(\frac{1}{N} \sum_{j=1}^N K(X_{i,t}^N - X_{j,t}^N) - V_{i,t}^N \right) dt + \sqrt{2} dB_{i,t}, \quad (1)$$

where the $(B_{i,t})_{t \geq 0}$ are independent Brownian motions. The interaction kernel is defined (everywhere) by

$$K(x) := \pm \frac{1}{2} \text{sign}(x) = \pm \frac{1}{2} \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x = 0, \\ -1 & \text{if } x < 0. \end{cases} \quad (2)$$

The case $K = \frac{1}{2} \text{sign}$ corresponds to the repulsive case, while the interaction is attractive when $K = -\frac{1}{2} \text{sign}$. This two cases lead of course to very different dynamics, but concerning the propagation of chaos in finite time the sign of K is not very relevant, so we will handle both cases in the same way.

*Université d'Aix-Marseille, CNRS, École Centrale, I2M, UMR 7373, 13453 Marseille, France.
maxime.hauray@univ-amu.fr, samir.salem@univ-amu.fr

Well-posedness of the particle system. The first non-obvious problem is raised by the particle system (1). Since the force field is only of bounded variation (BV in short) and not Lipschitz, the standard theory does not apply. Neither does the theory of existence and uniqueness developed for uniformly elliptic diffusions (see for instance [16, 5, 2] among many others) since the diffusion act here only on the velocities. However, due to the particular geometry of the problem, we can still get weak existence and uniqueness in law. The precise result is the following

Theorem 1. *For any $N \geq 2$, and any (deterministic) initial condition $(X_{i,0}^N, V_{i,0}^N)_{i \leq N} \in \mathbb{R}^{2N}$, weak existence and uniqueness in law hold for SDE (1).*

Since (1) is linear, it also implies weak existence and uniqueness for any random initial condition. Theorem 1 is proved in Section 2 using the following strategy: we reformulate (1) in an SDE with memory involving the $(V_{i,t}^N)_{i \leq N, t \geq 0}$ only (simply because the $X_{i,t}^N$ are time integrals of the $V_{i,t}^N$). Then we apply to that new (non-markovian) SDE a standard technique relying on the Girsanov's theorem. To the best of our knowledge, the strong existence and uniqueness of solution to that system are yet unknown, and we were not able to prove it.

A non-linear SDE related to the Vlasov-Poisson-Fokker-Planck equation. When the number of particles is large, and under the assumption that two particles picked up among all the others are roughly independent at any time (in particular this should be true at $t = 0$), we expect the N particles to behave almost like N i.i.d copies of the (expected unique) solution to the following non linear SDE, or McKean-Vlasov process:

$$dY_t = W_t dt, \quad dW_t = \mathbb{E}_{\bar{Y}_t} [K(Y_t - \bar{Y}_t)] dt - W_t dt + \sqrt{2} dB_t, \quad (3)$$

where $(B_t)_{t \geq 0}$ is a Brownian motion (independent of the rest) and \bar{Y}_t is an independent copy of Y_t .

We will prove the well-posedness of that nonlinear SDE for initial data with a uniform control on the velocity tails, and even a weak-strong stability estimate (here “strong solution” means that the law $\mathcal{L}(Y_t)$ of Y_t remains uniformly bounded in time). Before stating the precise result, we introduce two useful norms:

Definition 1. *For any $\lambda > 0$, and any $\gamma > 0$, we define for any $f \in L^\infty(\mathbb{R}^2)$ the two following norms:*

$$\|f\|_{e,\lambda} := \text{ess-sup}_{(x,v) \in \mathbb{R}^2} f(x,v) e^{\lambda|v|} \in [0, +\infty] \quad (4)$$

$$\|f\|_{p,\gamma} := \text{ess-sup}_{(x,v) \in \mathbb{R}^2} f(x,v) \langle v \rangle^{-\gamma} \in [0, +\infty] \quad (5)$$

where $\langle v \rangle = \sqrt{1+v^2}$, so that equivalently $\|f\|_{p,\gamma} = \text{ess-sup}_{(x,v) \in \mathbb{R}^2} f(x,v) (1+v^2)^{-\gamma/2}$, and the essential supremum are taken with respect to the Lebesgue measure.

Theorem 2. (i) *Existence of strong solutions:*

Assume that the law f_0 of the initial condition (Y_0, W_0) satisfies $f_0 \in \mathcal{P}_1(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)$ and also $\|f_0\|_{e,\lambda} < +\infty$, for some $\lambda > 0$ or $\|f_0\|_{p,\gamma} < +\infty$, for some $\gamma > 1$. Then, given any Brownian motion $(B_t)_{t \geq 0}$ (independent of the initial conditions) there exists a solution (Y_t, W_t) to (3) with initial condition (Y_0, W_0) , and its law at time $t \geq 0$: $f_t = \mathcal{L}(Y_t, W_t)$ satisfies either

$$\|f_t\|_{e,\lambda} \leq 2 e^{t+\lambda+\frac{\lambda^2}{2}} \|f_0\|_{e,\lambda e^{-t}}, \quad \text{or} \quad \|f_t\|_{p,\gamma} \leq C_\gamma e^{\gamma t} \|f_0\|_{p,\gamma}, \quad (6)$$

where C_γ is a constant depending only on γ .

(ii) *Weak/strong stability for solution with bounded density (in y).*

If (Y_t^1, W_t^1) and (Y_t^2, W_t^2) are two solutions to (3) built on the same probability space with the same Brownian motion $(B_t)_{t \in \mathbb{R}}$, if the density $\rho_t^1 = \mathcal{L}(Y_t^1)$ is uniformly bounded at any time: $\|\rho_t\|_\infty < +\infty$ for any $t \geq 0$, then the following stability estimate holds

$$\mathbb{E}[|Y_t^1 - Y_t^2| + |W_t^1 - W_t^2|] \leq e^{8(t + \int_0^t \|\rho_s^1\|_\infty ds)} \mathbb{E}[|Y_0^1 - Y_0^2| + |W_0^1 - W_0^2|]. \quad (7)$$

The proof of the weak-strong stability estimate relies on the crucial Lemma 4 that allows to control the singularity of the force, when comparing the evolution of two solutions, assuming only that one of them has a bounded density in position. That Lemma was already used in [10], to get similar results for the associated deterministic particle system: *i.e.* particles interacting via the same kernel but without noise and friction. The proof of the existence of solutions relies on an usual approximation procedure, see Section 3.3. The propagation of the bound on $\|f_t\|_{e,\lambda}$ or $\|f_t\|_{p,\gamma}$ is done using a standard argument that we may call the method of characteristics or Feynman-Kac's formula.

The stability result on the Vlasov-Poisson-Fokker-Planck equation The stability results on the process (3) simply translate on the associated Fokker-Planck or Kolmogorov forward equation, which is here the Vlasov-Poisson-Fokker-Planck equation (VPFP in short):

$$\partial_t f_t + v \partial_x f_t + (\rho_t \star K) \partial_v f_t = \partial_v (\partial_v f_t + v f_t), \quad (8)$$

where f_t is the law at time t of the process (Y_t, W_t) and $\rho_t = \int f_t dv$ is the law at time t of Y_t . As the kernel K is bounded and defined everywhere by (2), remark that very few hypothesis are required to define solutions to (8) in the sense of distribution: if $f \in L_{loc}^1(\mathbb{R}^+, \mathcal{P}(\mathbb{R}^2))$, where $\mathcal{P}(\mathbb{R}^2)$ stand for the space of probability, then all the terms appearing in (8) define a distribution.

The Theorem 2 as the following consequence on VPFP:

Corollary 1 (of Theorem 2). *(i) Existence of strong solution. Let $f_0 \in \mathcal{P} \cap L^1(\mathbb{R}^2)$ with a finite order one moment: $\int (|x| + |v|) f_0(dx, dv)$, and satisfying either $\|f_0\|_{e,\lambda} < +\infty$, for some $\lambda > 0$ or $\|f_0\|_{p,\gamma} < +\infty$, for some $\gamma > 1$. Then, there exists a solution f_t to (8) with initial condition f_0 , and it satisfies (6).*

(ii) Weak/strong uniqueness for solution with bounded density (in y).

If f_t^1 and f_t^2 are two solutions to (8) and if $\rho_t^1 = \int f_t^1 dv$ is uniformly bounded for any time $t \geq 0$, then the following stability estimate holds:

$$W_1(f_t^1, f_t^2) \leq e^{8(t + \int_0^t \|\rho_s^1\|_\infty ds)} W_1(f_0^1, f_0^2) \quad (9)$$

This corollary is proved in Section 3.4. It is a direct consequence of Theorem 2, and of the fact that a weak solution f_t to the VPFP1D equation (8), can always be represented as the time marginals of a process $(Y_t, W_t)_{t \geq 0}$ solution to (3).

The quantitative propagation of chaos in the mean When comparing a solution $(X_{i,t}^N, V_{i,t}^N)_{i \leq N}$ of the particle system (1) with the limit process (3) and its associated Fokker-Planck equation (8), a very natural strategy (which goes back to McKean [13]) is to introduce N independent copies $(Y_{i,t}^N, W_{i,t}^N)_{i \leq N}$ of the limit nonlinear SDE (3), constructed with the same Brownian motion as the $(X_{i,t}^N, V_{i,t}^N)$, and with initial conditions coupled in a optimal way. In our case, we are able to prove a sharp estimate on the average distance between these two systems.

To state our result properly, we recall that by definition exchangeable random variables have a law that is invariant under permutation, and that chaotic sequences of r.v. are defined as follows:

Definition 2. *Let f be a probability on \mathbb{R}^2 . A sequence $((X_i^N, V_i^N)_{i \leq N})_{N \in \mathbb{N}}$ of exchangeable random variables is said to be f -chaotic, if one of the equivalent conditions below is satisfied:*

- i) $\forall k \in \mathbb{N}, \quad \mathcal{L}((X_i^N, V_i^N)_{i \leq k}) \xrightarrow[N \rightarrow \infty]{w} f^{\otimes k},$
- ii) $\mathcal{L}((X_1^N, V_1^N), (X_2^N, V_2^N)) \xrightarrow[N \rightarrow \infty]{w} f \otimes f,$
- iii) $\mu^N := \frac{1}{N} \sum_{i=1}^N \delta_{(X_i^N, V_i^N)} \xrightarrow[N \rightarrow \infty]{\mathcal{L}} f.$

If Y is a random variable of law f , we will equivalently say that a sequence is f -chaotic or Y -chaotic. We refer to [14] for the equivalence of the three conditions above, and to [12] for a quantitative version of that equivalence.

We also state the following proposition, that reformulate propagation of chaos in term of coupling. It is a consequence of [12, Theorem 1.2]

Proposition 1. *Assume that $((X_i^N, V_i^N)_{i \leq N})_{N \in \mathbb{N}}$ is a sequence of exchangeable random variables with uniformly bounded order two moment: $\sup_{N \in \mathbb{N}} \mathbb{E}[|X_1^N|^2 + |V_1^N|^2] < +\infty$. Let f be a probability on \mathbb{R}^2 . Then, the following statements are equivalent:*

- i) *The sequence $((X_i^N, V_i^N)_{i \leq N})_{N \in \mathbb{N}}$ is f -chaotic;*
- ii) $\mathbb{E}\left[W_1(\mu^N, f)\right] \xrightarrow[N \rightarrow \infty]{} 0$, *where W_1 stands for the order one Monge-Kantorovich-Wasserstein distance, and $\mu^N = \frac{1}{N} \sum_{i=1}^N \delta_{(X_i^N, V_i^N)}$ is the associated empirical measure;*
- iii) *If the $(Y_i, W_i)_{i \in \mathbb{N}}$ are i.i.d.r.v. with common law f , independent of the $(X_i^N, V_i^N)_{i \leq N}$ for all N , then*

$$\mathbb{E}\left[|X_1^N - Y_1| + |V_1^N - W_1|\right] = \mathbb{E}\left[\frac{1}{N} \sum_i |X_i^N - Y_i| + |V_i^N - W_i|\right] \xrightarrow[N \rightarrow \infty]{} 0;$$

- iv) $\min_{\text{coupling}} \mathbb{E}\left[\frac{1}{N} \sum_i |X_i^N - Y_i| + |V_i^N - W_i|\right] \xrightarrow[N \rightarrow \infty]{} 0$, *where the minimum is taken on all the exchangeable coupling with (Y_i, V_i) i.i.d. with common law f .*

The above Proposition allows to state a precise result of propagation of molecular chaos. We emphasize that this result is a true result of propagation: it does not apply only to i.i.d. initial conditions, but to any chaotic initial conditions (with finite second order moment). However, the general case is somewhat more technical, so we warn the reader that the following theorem is simpler to understand if we only consider the case of i.i.d. initial conditions.

Theorem 3. *Let $f_0 \in \mathcal{P}(\mathbb{R}^2)$ with finite order two moment: $\int(|x|^2 + |v|^2) f_0(dx, dv) < \infty$, and such that there exists a (necessary unique by Corollary 1) solution f_t to (8) with initial condition f_0 satisfying $\int_0^t \|\rho_s\|_\infty ds < +\infty$ for any time $t \geq 0$, where ρ_s stands for the density in position: $\rho_s(x) := \int f_s(x, dv)$. We also denote by $(Y_t, W_t)_{t \geq 0}$ the unique solution to (3) such that $\mathcal{L}(Y_0, W_0) = f_0$.*

Let $(X_{i,0}^N, V_{i,0}^N)_{i \leq N}$ be a sequence of f_0 -chaotic random variable with uniformly bounded order two moment: $\sup_{N \in \mathbb{N}} \mathbb{E}[|X_{1,0}^N|^2 + |V_{1,0}^N|^2] < +\infty$. By Theorem 1, we may find a probability space together with a N -dimensional Brownian motion and a process $(X_{i,t}^N, V_{i,t}^N)_{i \leq N}$ solution to (1). Then, the sequence $((X_{i,t}^N, V_{i,t}^N)_{i \leq N})_{n \in \mathbb{N}}$ is $(Y_t, W_t)_{t \geq 0}$ chaotic.

More precisely, by standard arguments, we may also construct on that probability space N i.i.d. copies $(Y_{i,t}^N, W_{i,t}^N)_{i \leq N}$ of the solutions of (3) with the same Brownian motion, and with initial conditions of law f_0 , coupled with $(X_{i,0}^N, V_{i,0}^N)_{i \leq N}$ in an exchangeable way. Then the following estimate

holds:

$$\mathbb{E} \left[\sup_{s \in [0, t]} |X_{1,s}^N - Y_{1,s}^N| + |V_{1,s}^N - W_{1,s}^N| \right] \leq \left(\mathbb{E} [|X_{1,0}^N - Y_{1,0}^N| + |V_{1,0}^N - W_{1,0}^N|] + \frac{9}{\sqrt{N}} \right) e^{(t+8 \int_0^t \|\rho_s\|_\infty ds)}.$$

In particular, if $(X_{i,0}^N, V_{i,0}^N)_{i \leq N}$ are i.i.d. with law f_0 , then

$$\mathbb{E} \left[\sup_{s \in [0, t]} |X_{1,s}^N - Y_{1,s}^N| + |V_{1,s}^N - W_{1,s}^N| \right] \leq \frac{9}{\sqrt{N}} e^{(t+8 \int_0^t \|\rho_s\|_\infty ds)}.$$

Remark 1. Thanks to Corollary 1, the last hypothesis on f_0 is satisfied if $\|f_0\|_{e,\lambda} < \infty$ for some $\lambda > 0$, or if $\|f_0\|_{p,\gamma} < \infty$ for some $\gamma > 1$.

The proof of that result is performed in Section 4: it relies in a crucial way on a “rope argument” introduced in Section 3.1. A second key argument is the introduction of an ad-hoc Poisson Random Measure (PRM), which allows to conclude using standard properties of PRMs.

Using standard results on the convergence on empirical measures toward their mean [8, Theorem 1], we may deduce convergence results between the empirical measure of the particle system and the expected limit profile.

Corollary 2 (of Theorem 3). *Under the same assumptions than in Theorem 3, and assuming moreover that $\int (|x| + |v|)^q f_0(dx, dv) < \infty$ for some $q > 2$ we obtain that for any time $t \geq 0$,*

$$\mathbb{E}[W_1(\mu_t^N, f_t)] \leq C_t \left(\frac{\ln(1+N)}{\sqrt{N}} + \mathbb{E}[W_1(\mu_0^N, f_0)] \right),$$

for some constant $C_t > 0$ depending on t, q and f_0 .

Exponential concentration inequalities for the particle system. In the case where the initial conditions are i.i.d., we also prove concentration inequalities for the solutions of the particle system (1), precisely:

Theorem 4. *Let $f_0 \in \mathcal{P}(\mathbb{R}^2)$ with some finite exponential moment: $\int e^{\lambda(|x|+|v|)} f_0(dx, dv) < \infty$ for some $\lambda > 0$, and such that there exists a (necessary unique by Corollary 1) solution f_t to (8) with initial condition f_0 satisfying $\kappa_t := \sup_{s \leq t} \|\rho_s\|_\infty ds < +\infty$ for any time $t \geq 0$, where ρ_s stands for the density in position: $\rho_s(x) := \int f_s(x, dv)$. Let $(X_{i,0}^N, V_{i,0}^N)_{i \leq N}$ be a sequence of N i.i.d random variables with common law f_0 .*

By Theorem 1, we may find a probability space together with a N -dimensional Brownian motion and a process $(X_{i,t}^N, V_{i,t}^N)_{i \leq N}$ solution to (1). By standard arguments, we may also construct on that probability space N copies $(Y_{i,t}^N, W_{i,t}^N)_{i \leq N}$ of the solutions of (3) with the same initial conditions $(X_{i,0}^N, V_{i,0}^N)_{i \leq N}$ and Brownian motion.

Then the following concentration inequality holds for $\lambda N^{-1/2} \leq \varepsilon \leq (5\kappa_t \wedge 1) \min(\frac{1}{16}, \frac{\lambda}{2}, \lambda^{-2})$:

$$\mathbb{P} \left(\frac{1}{N} \sum_{i=1}^N \sup_{s \in [0, t]} (|X_{i,s}^N - Y_{i,s}^N| + |V_{i,s}^N - W_{i,s}^N|) \geq B_t \varepsilon \right) \leq (t + \varepsilon) (A_t + A'_t \sqrt{N} \varepsilon) N^{\frac{3}{2}} e^{-2N\varepsilon^2},$$

where the three constants depend on t, λ and the initial conditions f_0 . See the end of Section 6 for precise values.

Remark 2. Thanks to Corollary 1, the hypothesis on f_0 are satisfied if $\|f_0\|_{e,\lambda'} < \infty$ for some $\lambda' > \lambda$.

The proof of Theorem 4 relies on a different technique than the one used in the proof of Theorem 3. Here, we rather use exponential concentration inequalities on discrete infinite norms of empirical measures, and on some fluctuation terms appearing naturally when comparing solutions to (1) to copies of solutions to the nonlinear SDE (3).

Using deviation upper bounds for the approximation of probability measure by random empirical measures associated to i.i.d sample, as for instance in [8, Theorem 2], we can also obtain the following corollary.

Corollary 3. *Under the same assumption as in Theorem 4, and if moreover the initial positions and velocities $(X_{i,0}^N, V_{i,0}^N)$ are i.i.d random variables with law f_0 , then for any $T \geq 0$ and any $\lambda > 0$ there exists two constants C_1, C_2 such that for any $\varepsilon \geq 0$ satisfying $\lambda N^{-12} \leq \varepsilon \leq (5\kappa_t \wedge 1) \min(\frac{1}{16}, \lambda, \lambda^{-2})$, we have:*

$$\sup_{t \in [0, T]} \mathbb{P} \left[W_1(\mu_{X,t}^N, f_t) \geq \varepsilon \right] \leq C_1 \left(N^4 e^{-C_2 N \varepsilon^2} + e^{-C_2 (N \varepsilon)^{1-\alpha}} + e^{-C_2 N \varepsilon^2 / (1 - \ln \varepsilon)^2} \right),$$

where f_t is the unique solution of the VPFP equation (8) with initial condition f_0 .

Exponential concentration for discrete infinite norms. One of the key ingredient of the proof of Theorem 4 is a deviation inequality on time integral (or supremum) of discrete infinite norms for the $\mu_{Y,t}^N := N^{-1} \sum_i \delta_{Y_t^{N,i}}$, where the $Y_t^{N,i}$ are N i.i.d. r.v. solutions to a given SDE. Such a result have a interest by itself so we state it below

Proposition 2. *Let $(Y_i, W_i)_{i \in N}$ be N i.i.d copies of a solution of (3) driven by independent Brownian motions $(B_{i,t})_{i \leq N, t \geq 0}$, and assume that for $t \geq 0$:*

- *the common initial condition has a law f_0 which admits a finite exponential moment (in position and velocity) for some $\lambda > 0$: $\mathbb{E} \left[e^{\lambda(|Y_0| + |W_0|)} \right] < +\infty$. In particular we denote $c_\lambda := \frac{5}{2} + \frac{1}{\lambda} \ln \mathbb{E} \left[e^{\lambda|W_0|} \right]$;*
- $\kappa_t := \sup_{0 \leq s \leq t} \|\rho_s\|_\infty < +\infty$ *where ρ_s stands for the time marginal of Y_i at time s .*

Then provided that $\lambda N^{-1/2} \leq \varepsilon \gamma \leq 5\kappa_t \min(\frac{1}{16}, \lambda^{-2})$, the following bound holds with $\mathbf{C}_t := 10 + \kappa_t^2 \lambda^{-1} + e^{\lambda(1/2 + \lambda)t} \mathbb{E} \left[e^{\lambda(|Y_0| + |W_0|)} \right]$:

$$\mathbb{P} \left(\sup_{s \in [0, t]} \|\rho_s^N\|_{\infty, \varepsilon} \geq \kappa_t + \gamma \right) \leq \mathbf{C}_t \left(1 + t \frac{2\kappa_t + \gamma}{\lambda} (c_\lambda + \sqrt{N} \varepsilon \gamma) \right) N^{\frac{3}{2}} e^{-2N(\varepsilon \gamma)^2},$$

Essentially, the bound behave like $N^{3/2} e^{-2N(\varepsilon \gamma)^2}$, in the most interesting range, when $\varepsilon \gamma \sim N^{-1/2}$.

Some related works The literature on the convergence of particle systems towards non linear mean-field models is quite huge, so we will restrict ourselves to one dimensional models. The usual strategy, valid for smooth interaction, is well explained in Lecture notes by Sznitman [14]. In [3] Cépa and Lépine prove the propagation of chaos for the Dyson model: an order one model (*i.e.* without velocities) with a strongly singular interaction $K(x) \sim |x|^{-1}$ modeling the behavior of eigenvalues of large hermitian matrices. Their proof relies on the use of maximal monotone operators. Recently, that convergence result was extended to similar systems with even stronger interaction $K(x) \sim |x|^{-1-\alpha}$ with $\alpha \in [0, 1)$, by Berman and Önnheim [1] using Wasserstein gradient flows. For second order models (involving positions and velocities), our result is to the best of our

knowledge the first result of propagation of chaos with the Poisson singularity in the stochastic case. In the deterministic case, *i.e.* when the system under study is (1) without the Brownian motions, then the mean field limit was proved originally by [15], and then by [4] as a special case of semi-geostrophic equations, and again by the first author [10]. We shall finally mention the very recent preprint of Jabin and Wang [11], which proves the propagation of chaos in a very similar setting. They consider system like (1), with or without noise, with a very weak assumption on the interaction: K bounded, but a strong regularity assumption on the limit, and prove some quantitative propagation of chaos in term of relative entropy.

Plan of the paper. The paper is organized in the following way. In section 2, we prove Theorem 1 (weak existence and uniqueness of the particle system). In Section 3, we focus on the nonlinear limit SDE and prove Theorem 2 and its corollary 1, and also a useful proposition about propagation of moments (Proposition 3). Section 4 is devoted to the proof of Theorem 3 on the propagation of chaos in the mean, Section 5 to the proof of Proposition 2 and Section 6 to the proof of Theorem 4 on the exponential concentration. A useful regularity lemma is proved in the Appendix.

2 Proof of Theorem 1

A related SDE with memory. First note that a weak solution to equation (1) is some stochastic basis, together with a N dimensional Brownian motion $(B_t^N)_{t \geq 0}$ on it and a \mathbb{R}^{2N} valued processes $(X_t^N, V_t^N)_{t \geq 0}$ satisfying for all $t \geq 0$

$$X_t^N = X_0^N + \int_0^t V_s^N ds, \quad V_t^N = V_0^N + \int_0^t K^N(V_s^N) ds - \int_0^t V_s^N ds + \sqrt{2} B_t^N,$$

where we denoted $K^N(x_1, \dots, x_N)$ the vector valued field which i -th component is $\frac{1}{N} \sum_{j \neq i} K(x_i - x_j)$. But, from that system, we may write a SDE “with memory” involving V^N only:

$$\forall t \geq 0, \quad V_t^N = V_0^N + \int_0^t K^N \left(X_0^N + \int_0^s V_u^N du \right) ds - \int_0^t V_s^N ds + \sqrt{2} B_t^N. \quad (10)$$

Conversely, given a solution to the SDE (10), it is not difficult to construct a solution to the original system. So it will be enough to prove weak existence and uniqueness in law for the delayed SDE (10).

Weak existence. Let $(\Omega, \mathcal{F}, (\mathcal{F})_{t \geq 0}, \mathbb{P})$ be a stochastic basis and $(B_t^N)_{t \geq 0}$ be a N -dimensional on it. We define

$$U_t^N := -\frac{1}{\sqrt{2}} \int_0^t K^N \left(X_0^N + s V_0^N + \sqrt{2} \int_0^s B_u^N du \right) ds + \int_0^t \left(B_s^N + \frac{1}{\sqrt{2}} V_0^N \right) ds + B_t^N, \quad V_t^N := V_0^N + \sqrt{2} B_t^N.$$

The above definition of the two r.v. $(U_t^N, V_t^N)_{t \geq 0}$ implies that for any $t \geq 0$,

$$V_t^N = V_0^N + \int_0^t K^N \left(X_0^N + \int_0^s V_u^N du \right) ds - \int_0^t V_s^N ds + \sqrt{2} U_t^N,$$

which is exactly (10) with B_t^N replaced by U_t^N . So, it remains to apply Cameron-Martin-Girsanov (CMG) theorem: with an appropriate change of the reference probability measure, $(U_t^N)_{t \geq 0}$ can be considered has a N -dimensional Brownian motion. For this, remark that $dU_t^N = -H_t^N dt + dB_t^N$, where

$$H_t^N := \frac{1}{\sqrt{2}} K^N \left(X_0^N + t V_0^N + \int_0^t \sqrt{2} B_u^N du \right) - B_t^N - \frac{1}{\sqrt{2}} V_0^N,$$

is \mathcal{F}_t -adapted and progressively measurable, and that for $0 < \gamma < \frac{1}{6t}$

$$\mathbb{E} \left[e^{\gamma |H_t^N|^2} \right] \leq e^{\frac{3}{2}\gamma (N\|K\|_\infty^2 + |V_0^N|^2)} \mathbb{E} \left[e^{3\gamma (B_t^N)^2} \right] < +\infty.$$

Therefore we deduce from classical results about exponential martingales that the process Z_t^N defined by

$$Z_t^N = \exp \left(\int_0^t H_s^N \cdot dB_s^N - \frac{1}{2} \int_0^t |H_s^N|^2 ds \right)$$

(where \cdot stands for the scalar product) is a martingale, and due to CMG theorem $(U_t^N)_{t \geq 0}$ is a N -dimensional Brownian motion under the probability \mathbb{Q} defined for any A in \mathcal{F}_t by $\mathbb{Q}(A) = \int_A Z_t^N d\mathbb{P}$. Therefore $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{Q}, (U_t^N, V_t^N)_{t \geq 0})$ is a weak solution to SDE (10).

Uniqueness in law. Suppose that $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}, (V_t^N, B_t^N)_{t \geq 0})$ is a solution to equation (10) with initial condition V_0^N . We may use CMG theorem again: in fact $2^{-1/2}dV_t^N = -\tilde{H}_t^N dt + dB_t^N$, where

$$\tilde{H}_t^N := -\frac{1}{\sqrt{2}}K^N \left(X_0^N + \int_0^t V_u^N du \right) + \frac{1}{\sqrt{2}}V_t^N.$$

Moreover by (10), V_t^N also satisfies

$$V_t^N = e^{-t}V_0^N + \int_0^t e^{s-t}K^N \left(X_0^N + \int_0^s V_u^N du \right) ds + \sqrt{2} \int_0^t e^{s-t}dB_s^N,$$

which implies that

$$|\tilde{H}_t^N| \leq |V_t^N| + \frac{\sqrt{N}}{2} \leq |V_0^N| + \frac{3\sqrt{N}}{2} + |M_t^N|,$$

where M_t^N is a Gaussian r.v. with law $\mathcal{N}(0, (1 - e^{2t})\text{Id})$. It follows that $\mathbb{E}[e^{\gamma |\tilde{H}_t^N|^2}] < \infty$, for $t \geq 0$ and $\gamma < 1/2$. This implies that the process \tilde{Z}_t^N defined by

$$\tilde{Z}_t^N = \exp \left(\int_0^t \tilde{H}_s^N \cdot dV_s^N - \frac{1}{2} \int_0^t |\tilde{H}_s^N|^2 ds \right) = \exp \left(\frac{1}{\sqrt{2}} \int_0^t \tilde{H}_s^N \cdot dB_s^N + \frac{1}{2} \int_0^t |\tilde{H}_s^N|^2 ds \right),$$

is a martingale, and by Cameron-Martin-Girsanov theorem, $2^{-1/2}(V_t^N - V_0^N)_{t \geq 0}$ is a N -dimensional Brownian motion on the filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \tilde{\mathbb{Q}})$ where $\tilde{\mathbb{Q}}$ is defined for any A in \mathcal{F}_t by $\tilde{\mathbb{Q}}(A) = \int_A \tilde{Z}_t^N d\mathbb{P}$.

Now for any “cylindrical” function ϕ on $C(\mathbb{R}^+, \mathbb{R}^N)$ of the form $\phi((V_s^N)_{s \geq 0}) = \varphi_1(V_{t_1}^N) \times \cdots \times \varphi_k(V_{t_k}^N)$, we get for $t \geq t_k$:

$$\begin{aligned} \mathbb{E}_{\mathbb{P}} \left[\phi((V_s^N)_{s \geq 0}) \right] &= \mathbb{E}_{\tilde{\mathbb{Q}}} \left[\phi((V_s^N)_{s \geq 0}) (\tilde{Z}_t^N)^{-1} \right] \\ &= \mathbb{E}_{\tilde{\mathbb{Q}}} \left[\phi((V_s^N)_{s \geq 0}) \exp \left(-\frac{1}{\sqrt{2}} \int_0^t \tilde{H}_s^N \cdot dV_s^N - \frac{1}{2} \int_0^t |\tilde{H}_s^N|^2 ds \right) \right]. \end{aligned}$$

The expression on the last line does not involve $(B_t^N)_{t \geq 0}$ anymore. Since the law of $(V_t^N)_{t \geq 0}$ under $\tilde{\mathbb{Q}}$ is the law of the Brownian motion, and since \tilde{Z}_t^N can be expressed in term of V_t^N only, that last expression does not depend on the specific solution we selected at the beginning of this paragraph: we will obtain exactly the same formula starting from a second solution. Since, solution to (10) have continuous trajectories, this implies the uniqueness in law of the solutions to the SDE (10) and also to (1).

3 Proof of Theorem 2

3.1 Weak-strong stability and uniqueness.

A key bound. Here we will use a “rope” argument which has already been used in [10] to treat the propagation of chaos for deterministic VP1D equation. It consists in noticing that:

$$K(x - \bar{x}) - K(y - \bar{y}) = 0$$

as soon as $|y - \bar{y}| > |x - y| + |\bar{x} - \bar{y}|$. It is also interesting to replace the later condition by the stronger one $|y - \bar{y}| > 2 \max(|x - y|, |\bar{x} - \bar{y}|)$. Since K is bounded by $1/2$, it implies the following bound, that we will use many times in the sequel

$$|K(x - \bar{x}) - K(y - \bar{y})| \leq \mathbb{I}_{|y - \bar{y}| \leq 2|x - y|} + \mathbb{I}_{|y - \bar{y}| \leq 2|\bar{x} - \bar{y}|} \quad (11)$$

That one-sided condition (the indicator functions in the r.h.s. take only $y - \bar{y}$ as argument) allows to prove a simple but important Lemma, where discrete infinite norms are introduced:

Definition 3. For any $\varepsilon > 0$, and any $f \in \mathcal{P}(\mathbb{R})$, we define the infinite norm at scale ε , denoted $\|f\|_{\infty, \varepsilon}$ by:

$$\|f\|_{\infty, \varepsilon} := \sup_{x \in \mathbb{R}} \frac{f([x - \varepsilon, x + \varepsilon])}{2\varepsilon} = \left\| f \star \frac{1}{2\varepsilon} \mathbb{I}_{[-\varepsilon, \varepsilon]} \right\|_{\infty}.$$

Lemma 4. Assume that (X, Y) is a random couple of real numbers and that (\bar{X}, \bar{Y}) is an independent copy of that couple.

(i) Then,

$$\mathbb{E}[K(X - \bar{X}) - K(Y - \bar{Y})] \leq 8 \min(\|\rho_X\|_{\infty}, \|\rho_Y\|_{\infty}) \mathbb{E}[|X - Y|],$$

where ρ_X, ρ_Y denote respectively the density (with respect to the Lebesgue measure) of the law of X and Y .

(ii) For $\varepsilon > 0$, we also have a similar estimate involving the discrete infinite norms $\|\cdot\|_{\infty, \varepsilon}$ defined in Definition 3:

$$\mathbb{E}[|K(X - \bar{X}) - K(Y - \bar{Y})|] \leq 8 \min(\|\rho_X\|_{\infty, \varepsilon}, \|\rho_Y\|_{\infty, \varepsilon}) \left(\mathbb{E}[|X - Y|] + \frac{\varepsilon}{2} \right).$$

(iii) In the case where (X, Y) and (\bar{X}, \bar{Y}) are still independent but with possibly different distributions, we get the more general estimate: for $\varepsilon, \bar{\varepsilon} \geq 0$ (in case set $\varepsilon = 0$ or $\bar{\varepsilon} = 0$, set $\|\cdot\|_{\infty, 0} = \|\cdot\|_{\infty}$),

$$\mathbb{E}[|K(X - \bar{X}) - K(Y - \bar{Y})|] \leq 4 \|\rho_{\bar{Y}}\|_{\infty, \bar{\varepsilon}} \left(\mathbb{E}[|X - Y|] + \bar{\varepsilon}/2 \right) + 4 \|\rho_Y\|_{\infty, \varepsilon} \left(\mathbb{E}[|\bar{X} - \bar{Y}|] + \varepsilon/2 \right).$$

Proof. We first prove i). Starting from (11), we may bound

$$\begin{aligned} \mathbb{E}[|K(X - \bar{X}) - K(Y - \bar{Y})|] &\leq \mathbb{E}[\mathbb{I}_{|Y - \bar{Y}| \leq 2|X - Y|} + \mathbb{I}_{|Y - \bar{Y}| \leq 2|\bar{X} - \bar{Y}|}] \\ &\leq 2 \mathbb{E}[\mathbb{I}_{|Y - \bar{Y}| \leq 2|X - Y|}] \\ &\leq 2 \mathbb{E}[\mathbb{E}[\mathbb{I}_{|Y - \bar{Y}| \leq 2|X - Y|} | (X, Y)]] \\ &\leq 2 \mathbb{E}[4 \|\rho_Y\|_{\infty} |X - Y|] \\ &\leq 8 \|\rho_Y\|_{\infty} \mathbb{E}[|X - Y|]. \end{aligned}$$

In the second line, we used that the expectation remain unchanged if we permute (X, Y) with (\bar{X}, \bar{Y}) . In the fourth, we used that the law of \bar{Y} has density ρ_Y . If we apply the previous calculation with the couple (Y, X) and (\bar{Y}, \bar{X}) , we obtain a similar result, with $\|\rho_X\|_\infty$ in place of $\|\rho_Y\|_\infty$.

To obtain *ii*), remark that for any interval $[a, b] \subset \mathbb{R}$:

$$\mathbb{E}[\mathbb{1}_A(X)] = \int_a^b \rho_X(dx) \leq \|\rho_X\|_{\infty, \varepsilon}(|b - a| + 2\varepsilon).$$

Indeed, to use discrete infinite norm, we need to cover $[a, b]$ by a union of small intervals of length 2ε . For this at most (the integer part of) $|b - a|/(2\varepsilon) + 1$ such intervals are requested.

For the third point, we cannot use the permutation $(X, Y) \leftrightarrow (\bar{X}, \bar{Y})$ in the previous calculation and have to estimate the two terms separately. The necessary adaptations are straightforward. \square

A simple Grönwall lemma. That bound allows us to prove the weak-strong stability part of Theorem 2. We introduce $(X_t, V_t)_{t \in \mathbb{R}_+}$ and $(Y_t, W_t)_{t \in \mathbb{R}_+}$ two solutions of the non-linear SDE (3) constructed on the same Brownian motion $(B_t)_{t \in \mathbb{R}}$, and also $(\bar{X}_t, \bar{V}_t, \bar{Y}_t, \bar{W}_t)_{t \in \mathbb{R}}$ an independent copy of the previous coupled processes. We also assume that $\rho_t := \mathcal{L}(X_t)$ has a bounded density for all $t \geq 0$. Then $(X_t - Y_t, V_t - W_t)_{t \in \mathbb{R}}$ solves the following ODE system:

$$\frac{d}{dt}(X_t - Y_t) = V_t - W_t, \quad \frac{d}{dt}(V_t - W_t) = -(V_t - W_t) + \mathbb{E}_{(\bar{X}_t, \bar{Y}_t)}[K(X_t - \bar{X}_t) - K(Y_t - \bar{Y}_t)],$$

which naturally leads to

$$\begin{aligned} \sup_{s \in [0, t]} |X_s - Y_s| &\leq |X_0 - Y_0| + \int_0^t |V_s - W_s| ds, \\ \sup_{s \in [0, t]} |V_s - W_s| &\leq |V_0 - W_0| + \int_0^t |K(X_s - \bar{X}_s) - K(Y_s - \bar{Y}_s)| ds. \end{aligned}$$

If we take the expectation in the previous system, and apply the point (i) of Lemma 4 to the couple (Y_t, X_t) and (\bar{Y}_t, \bar{X}_t) , we may write:

$$\begin{aligned} \mathbb{E} \left[\sup_{s \in [0, t]} |X_s - Y_s| \right] &\leq \mathbb{E}[|X_0 - Y_0|] + \int_0^t \mathbb{E}[|V_s - W_s|] ds \\ \mathbb{E} \left[\sup_{s \in [0, t]} |V_s - W_s| \right] &\leq \mathbb{E}[|V_0 - W_0|] + 8 \int_0^t \|\rho_s\|_\infty \mathbb{E}[|X_s - Y_s|] ds. \end{aligned}$$

Summing up the two inequalities and applying the Grönwall lemma lead to the requested estimate

$$\mathbb{E} \left[\sup_{s \in [0, t]} (|X_s - Y_s| + |V_s - W_s|) \right] \leq \mathbb{E}[|X_0 - Y_0| + |V_0 - W_0|] \exp \left(t + 8 \int_0^t \|\rho_s\|_\infty ds \right).$$

Remark that it is not completely straightforward to take advantage of the restoring force in order to improve the above bound, especially because of the supremum in time.

3.2 Propagation of moments

Here, we show that order one moments, and exponential moments are propagated by the SDE (3). We emphasize that our results apply not only for K given by (2), but as soon as $\|K\|_\infty \leq 1/2$. In particular, we will apply it later to nonlinear SDE where K is replaced by a smooth mollification.

Proposition 3. Let be $(Y_t, W_t)_{t \geq 0}$ be a weak solution to 3, with given (random) initial condition (Y_0, W_0) , and with an interaction kernel K which is not necessary given by (2) but satisfies $\|K\|_\infty \leq \frac{1}{2}$. If (Y_0, W_0) has an exponential moment of order $\lambda > 0$, it holds for $t \geq 0$:

$$(i) \quad \mathbb{E}[e^{\lambda|W_t|}] \leq e^{\frac{\lambda}{2}(3+\lambda)} \mathbb{E}[e^{\lambda e^{-t}|W_0|}],$$

$$(ii) \quad \mathbb{E}[e^{\lambda|Y_t|}] \leq 2 e^{\lambda t(\frac{1}{2}+\lambda)} \mathbb{E}[e^{\lambda(|Y_0|+|W_0|)}] := 2 e^{\lambda t(\frac{1}{2}+\lambda)} \mathcal{M}_\lambda^{x,v}(f_0).$$

It also holds for any $0 \leq s < t \leq s + \min(\frac{1}{16}, \lambda^{-2})$:

$$(iii) \quad \mathbb{E}[e^{\lambda(t-s)^{-1} \sup_{s \leq u \leq t} |Y_u - Y_s|}] \leq e^{\frac{\lambda}{2}(5+\lambda)} \mathbb{E}[e^{\lambda|W_0|}].$$

Lastly, simpler estimates on the order one moments also hold: for any $0 \leq s < t \leq s + \frac{1}{4}$,

$$(iv) \quad \mathbb{E}[|W_t|] \leq e^{-t} \mathbb{E}[|W_0|] + 2, \quad \mathbb{E}[|Y_t - Y_s|] \leq |s - t|(\mathbb{E}[|W_0|] + 3).$$

Proof. Point (i). First, introducing the notation f_t for the time marginal of (Y_t, W_t) and $F(t, x) := \int K(x - y) f_t(dy, dw)$ we have by (3)

$$e^t W_t = W_0 + \int_0^t e^s F(s, Y_s) ds + \sqrt{2} \int_0^t e^s dB_s. \quad (12)$$

Next, since $|K|$ and also $|F|$ are bounded by $1/2$, we get a simple inequality

$$|W_t| \leq e^{-t}|W_0| + \frac{1}{2} + |M_t|, \quad \text{with} \quad M_t := \sqrt{2} \int_0^t e^{s-t} dB_s. \quad (13)$$

M_t is a centered Gaussian random variable with variance $1 - e^{-2t} \leq 1$, for which exponential moments are simple to obtain. In fact for a Gaussian variable $Z \sim \mathcal{N}(0, \sigma^2)$ we have a simple bound

$$E(e^{\lambda|Z|}) = \frac{2}{\sqrt{2\pi}\sigma} \int_0^{+\infty} e^{\lambda x - \frac{x^2}{2\sigma^2}} dx = \frac{2}{\sqrt{\pi}} e^{\frac{\lambda^2 \sigma^2}{2}} \int_{-\lambda\sigma/2}^{+\infty} e^{-x^2} dx = e^{\frac{\lambda^2 \sigma^2}{2}} \left(1 + \operatorname{erf}\left(\frac{\lambda\sigma}{\sqrt{2}}\right)\right),$$

where we used for the error function erf the definition $\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-x^2} dx$. Using that $\operatorname{erf}(x) \leq \min(1, e^x)$, we finally get the following bound, that will be very useful in the sequel:

$$E(e^{\lambda|Z|}) \leq \min\left(2, e^{\frac{1}{\sqrt{2}}\lambda\sigma}\right) e^{\frac{\lambda^2 \sigma^2}{2}} \quad (14)$$

Together with the independence of the Brownian motion $(B_t)_{t \geq 0}$ and the initial condition W_0 , it leads here to

$$\begin{aligned} \mathbb{E}[e^{\lambda|W_t|}] &\leq e^{\frac{\lambda}{2}} \mathbb{E}[e^{\lambda e^{-t}|W_0|}] \mathbb{E}[e^{\lambda|M_t|}] \leq e^{\frac{\lambda}{2}} \mathbb{E}[e^{\lambda e^{-t}|W_0|}] e^{\frac{1}{\sqrt{2}}\lambda + \frac{1}{2}\lambda^2} \\ &\leq e^{\frac{\lambda}{2}(3+\lambda)} \mathbb{E}[e^{\lambda e^{-t}|W_0|}], \end{aligned}$$

which is exactly (i).

Point (ii). For the second point, we integrate the inequality (12) and get:

$$Y_t = Y_0 + \int_0^t W_s ds = Y_0 + (1 - e^{-t})W_0 + \int_0^t (1 - e^{s-t})F(s, Y_s) ds + N_t, \quad (15)$$

$$|Y_t| \leq |Y_0| + |W_0| + \frac{t}{2} + |N_t|, \quad (16)$$

$$\text{with} \quad N_t := \int_0^t e^{-s} M_s ds = \sqrt{2} \int_0^t (1 - e^{u-t}) dB_u,$$

where we have used a stochastic version of Fubini's Theorem in the last line. N_t is a centered random Gaussian variable with variance $\sigma_t^2 := 4e^{-t} - e^{-2t} - 3 + 2t \leq 2t$. The bound (14) and the independence of the initial condition and N_t then lead to

$$\begin{aligned}\mathbb{E}\left[e^{\lambda|Y_t|}\right] &\leq e^{\lambda\frac{t}{2}}\mathbb{E}\left[e^{\lambda(|Y_0|+|W_0|)}\right]\mathbb{E}\left[e^{\lambda|N_t|}\right], \\ &\leq e^{\lambda\frac{t}{2}}\mathbb{E}\left[e^{\lambda(|Y_0|+|W_0|)}\right] 2e^{\frac{1}{2}\lambda^2\sigma_t^2} \leq 2e^{\lambda\frac{t}{2}+\lambda^2t}\mathbb{E}\left[e^{\lambda(|Y_0|+|W_0|)}\right],\end{aligned}$$

which leads to the claimed result.

Point (iii). We first integrate equality (12) between s and u , and take a supremum in time:

$$\begin{aligned}Y_u &= Y_s + (1 - e^{s-u})W_s + \int_s^u (1 - e^{v-u})F(v, Y_v) dv + N'_u, \\ \sup_{s \leq u \leq t} |Y_u - Y_s| &\leq (t-s)|W_s| + \frac{t-s}{2} + \sup_{s \leq u \leq t} |N'_u|, \quad \text{with } N'_u := \sqrt{2} \int_s^u (1 - e^{v-u}) dB_v, \quad (17)\end{aligned}$$

But, thanks to the properties of Brownian motion, by a change of time $\tau = \sigma_{t-s}^2 = 4e^{s-t} - e^{2(s-t)} - 3 + 2(t-s) \leq (t-s)^3$:

$$\sup_{s \leq u \leq t} |N'_u| \stackrel{\mathcal{L}}{=} \sup_{0 \leq u \leq \tau} |B_u| \stackrel{\mathcal{L}}{=} \sqrt{\tau} \sup_{0 \leq u \leq 1} |B_u|.$$

The law of supremum in time of the absolute value of a 1D Brownian motion is explicitly known, see for instance [6, p. 342]. Here, we will use only simple estimates on the exponential moments:

$$\begin{aligned}\mathbb{E}\left[e^{\lambda \sup_{0 \leq u \leq 1} |B_u|}\right] &\leq \mathbb{E}\left[e^{\lambda \sup_{0 \leq u \leq 1} B_u}\right] + \mathbb{E}\left[e^{\lambda \sup_{0 \leq u \leq 1} (-B_u)}\right] \\ &\leq 2\mathbb{E}\left[e^{\lambda |B_1|}\right] \leq 4e^{\frac{1}{2}\lambda^2}.\end{aligned}$$

In the second line, we use the well-known equality $\sup_{0 \leq u \leq \tau} B_u \stackrel{\mathcal{L}}{=} \sup_{0 \leq u \leq \tau} (-B_u) \stackrel{\mathcal{L}}{=} |B_\tau|$, and then the exponential moments given by (14). The constant 4 appearing above will raise some difficulties so we will perform a little optimization to get rid of it. For any $\theta \geq 1$, we may also bound

$$\mathbb{E}\left[e^{\lambda \sup_{0 \leq u \leq 1} |B_u|}\right] \leq \mathbb{E}\left[e^{\lambda \theta \sup_{0 \leq u \leq 1} |B_u|}\right]^{\frac{1}{\theta}} \leq \left(4e^{\frac{1}{2}(\theta\lambda)^2}\right)^{\frac{1}{\theta}} = e^{\frac{1}{\theta} \ln 4 + \frac{\theta}{2}\lambda^2}$$

The optimal θ seems to be $\theta = \frac{2\sqrt{\ln 2}}{\lambda}$. It is admissible when $\lambda \leq 2\sqrt{\ln 2}$. It leads to

$$\mathbb{E}\left[e^{\lambda \sup_{0 \leq u \leq 1} |B_u|}\right] \leq e^{2\lambda\sqrt{\ln 2}} \leq e^{2\lambda}.$$

Finally for $t-s \leq \lambda^{-2}$, the upper bound on τ leads to $\lambda(t-s)^{-1}\sqrt{\tau} \leq \lambda\sqrt{t-s} \leq 1$ and

$$\mathbb{E}\left[e^{\lambda(t-s)^{-1} \sup_{s \leq u \leq t} |N'_u|}\right] = \mathbb{E}\left[e^{\lambda(t-s)^{-1} \sqrt{\tau} \sup_{0 \leq u \leq 1} |B_u|}\right] \leq e^{2\lambda\sqrt{t-s}}.$$

Using the point (ii) of the Proposition and (17), we write

$$\begin{aligned}\mathbb{E}\left[e^{\lambda(t-s)^{-1} \sup_{s \leq u \leq t} |Y_u - Y_s|}\right] &\leq e^{\frac{\lambda}{2}}\mathbb{E}\left[e^{\lambda|W_s|}\right]\mathbb{E}\left[e^{\lambda(t-s)^{-1} \sup_{s \leq u \leq t} |N'_u|}\right] \\ &\leq e^{\frac{\lambda}{2}} e^{\frac{\lambda}{2}(3+\lambda)}\mathbb{E}\left[e^{\lambda|W_0|}\right] e^{2\lambda\sqrt{t-s}},\end{aligned}$$

which concludes the proof, using that $\sqrt{t-s} \leq \frac{1}{4}$ by assumption.

Point (iv). Taking the expectation in (13),

$$\mathbb{E}[|W_t|] \leq e^{-t}\mathbb{E}[|W_0|] + \frac{1}{2} + \mathbb{E}[|M_t|].$$

and using that M_t is $\mathcal{N}(0, 1 - e^{-2t})$ distributed, the expectation is simply bounded: $\mathbb{E}[|M_t|] \leq \sqrt{2/\pi} \leq 1$ and it implies the first bound of point (iv). The second bound uses (16) written with s instead of 0, which leads in average to

$$\mathbb{E}[|Y_t - Y_s|] \leq (1 - e^{s-t})\mathbb{E}[|W_s|] + \frac{t-s}{2} + \mathbb{E}[|N_{t-s}|].$$

Using, $1 - e^{s-t} \leq t - s$ and the previous bound on $\mathbb{E}[|W_s|]$, and the fact that N_t is $\mathcal{N}(0, \sigma_{t-s}^2)$ distributed, with $\sigma_{t-s}^2 \leq (t-s)^3$, it comes

$$\mathbb{E}[|Y_t - Y_0|] \leq (t-s) \left(\mathbb{E}[|W_0|] + 2 + \frac{1}{2} + \sqrt{\frac{2(t-s)}{\pi}} \right),$$

and the conclusion follows when $t - s \leq \frac{1}{4}$. \square

3.3 Strong existence via regularization and Feymann-Kac type estimates.

We introduce a smoothing kernel $\chi \in C^\infty(\mathbb{R}, \mathbb{R}^+)$ with support included in $[-1, 1]$ and satisfying $\int_{\mathbb{R}} \chi(y) dy = 1$. And then standardly for $\eta > 0$, $\chi_\eta := \eta^{-1} \chi(\frac{\cdot}{\eta})$, and the approximated kernel

$$K_\eta := K * \chi_\eta, \quad \text{which satisfies } |(K_\eta - K)(x)| \leq \mathbb{I}_{[-\eta, \eta]}(x), \quad (18)$$

for all $x \in \mathbb{R}$. Given a stochastic basis, and a Brownian $(B_t)_{t \geq 0}$ motion on it, we consider the following non linear SDE:

$$Y_t^\eta = Y_0^\eta + \int_0^t W_s^\eta ds, \quad W_t^\eta = W_0^\eta + \int_0^t \mathbb{E}_{\bar{Y}}[K_\eta(Y_s^\eta - \bar{Y})] ds - \int_0^t W_s^\eta ds + \sqrt{2} B_t, \quad (19)$$

where \bar{Y} is an independent copy of Y^η and the initial condition (Y_0^η, W_0^η) is defined as

$$Y_0^\eta := Y_0 + \eta U, \quad W_0^\eta := W_0 + \eta V, \quad (20)$$

where (Y_0, W_0) has law f_0 and is independent of (U, V) of law $\chi \otimes \chi$ (and both are independent of the Brownian motion $(B_t)_{t \geq 0}$). Then (Y_0^η, W_0^η) has for law $\mu_0^\eta := f_0 * \tilde{\chi}_\eta$ with $\tilde{\chi}_\eta(y, w) := \chi_\eta(y) \chi_\eta(w)$, and is independent of the Brownian motion $(B_t)_{t \geq 0}$. Introducing the notation

$$\tilde{K}_\eta[\mu](x) = \int_{\mathbb{R}^2} K_\eta(x - y) \mu(dy, dw),$$

this system can be written in an equivalent manner:

$$Y_t^\eta = Y_0^\eta + \int_0^t W_s^\eta ds, \quad W_t^\eta = W_0^\eta + \int_0^t \tilde{K}_\eta[\mu_s^\eta](Y_s^\eta) ds - \int_0^t W_s^\eta ds + \sqrt{2} B_t,$$

where μ_t^η is the time marginal at time t , i.e. the law of (Y_t^η, W_t^η) .

Since the kernel K_η is globally Lipschitz, [14, Thm 1.1] implies the strong existence and uniqueness of the process $(Y_t^\eta, W_t^\eta)_{t \geq 0}$ solving (19). And by an application of Ito's rule the family of the time marginals $(\mu_t^\eta)_{t \geq 0}$ of that process is a weak solution of the following regularized Vlasov-Poisson-Fokker-Planck equation:

$$\frac{\partial}{\partial t} \mu_t^\eta + v \partial_x \mu_t^\eta + \tilde{K}_\eta[\mu_t^\eta] \partial_v \mu_t^\eta = \partial_v (\partial_v \mu_t^\eta + v \mu_t^\eta), \quad (21)$$

with the initial condition $\mu_0^\eta = \mathcal{L}(Y_0, W_0) = f_0 * \tilde{\chi}_\eta$. We begin by proving some η independent estimates on μ_t^η , for $t \geq 0$.

Feynmann-Kac type estimates

Lemma 5. *Assume that the law of the initial condition of equation (3) $f_0 \in \mathcal{P}_1 \cap L^1(\mathbb{R}^2)$ satisfies either $\|f_0\|_{e,\lambda} < \infty$ for some $\lambda > 0$ or $\|f_0\|_{p,\gamma} < \infty$ for some $\gamma > 1$. Then for all $t > 0$, the unique (measure) solution to the smoothed VFPF equation (21) with initial condition μ_0^η satisfies respectively*

$$\|\mu_t^\eta\|_{e,\lambda} \leq 2e^{t+\lambda\eta+\frac{\lambda}{2}+\frac{\lambda^2}{2}} \|f_0\|_{e,\lambda e^{-t}}, \quad \|\mu_t^\eta\|_{p,\gamma} \leq \|f_0\|_{p,\gamma} \langle \eta \rangle^\gamma C_\gamma e^{(1+\gamma)t}, \quad (22)$$

where C_γ is a constant depending explicitly on γ .

In particular, the associated spatial density $\rho_t^\eta := \int_{\mathbb{R}} \mu_t^\eta(x, v) dv$ satisfies respectively

$$\|\rho_t^\eta\|_\infty \leq \frac{4}{\lambda} e^{t+\frac{\lambda}{2}+\frac{\lambda^2}{2}} e^{\lambda\eta} \|f_0\|_{e,\lambda e^{-t}}, \quad \|\rho_t^\eta\|_\infty \leq \frac{2\gamma}{\gamma-1} C_\gamma e^{(1+\gamma)t} \langle \eta \rangle^\gamma \|f_0\|_{p,\gamma}. \quad (23)$$

Proof. Step 1. Regularization and Feynmann-Kac's formula. Fix $t \geq 0$ and consider the following “backward” SDE:

$$Y_s^{x,v} = x - \int_0^s W_u^{x,v} du, \quad W_s^{x,v} = v - \int_0^s \tilde{K}_\eta[\mu_{t-u}^\eta](Y_u^{x,v}) du + \int_0^s W_u^{x,v} du + \sqrt{2} B_s, \quad (24)$$

First note that $\tilde{K}_\eta[\mu^\eta]$ is uniformly Lipschitz in position on $\mathbb{R}^+ \times \mathbb{R}^2$. So strong existence and uniqueness of solution to the (linear) SDE (24) are guaranteed by standard results. We set:

$$\theta_s = e^s \mu_{t-s}^\eta(Y_s^{x,v}, W_s^{x,v}).$$

Moreover the initial condition $\mu_0^\eta = f_0 * \tilde{\chi}_\eta$ fulfills the hypothesis of Proposition 4 of the Appendix: $\partial_x^k \partial_v^l \mu_0^\eta \in L^2(\mathbb{R}^2)$ for any $k, l \geq 0$. This implies that $\mu_t^\eta(x, v)$ possesses one continuous derivative in time, and two (continuous) derivative in position and velocity. So, we may apply Ito's rule to θ : we get

$$\begin{aligned} e^{-s} d\theta_s &= \mu^\eta(t-s, Y_s^{x,v}, W_s^{x,v}) ds - \partial_t \mu^\eta(t-s, Y_s^{x,v}, W_s^{x,v}) ds + \partial_x \mu^\eta(t-s, Y_s^{x,v}, W_s^{x,v}) dY_s^{x,v} \\ &\quad + \partial_v \mu^\eta(t-s, Y_s^{x,v}, W_s^{x,v}) dW_s^{x,v} + \Delta_v \mu^\eta(t-s, Y_s^{x,v}, W_s^{x,v}) \langle dW_s^{x,v} \rangle^2 \\ &= \left[-\partial_t \mu^\eta - v \partial_x \mu^\eta - \tilde{K}_\eta[\mu^\eta] \partial_v \mu^\eta + \partial_v(v \mu^\eta) + \Delta_v \mu^\eta \right] (t-s, Y_s^{x,v}, W_s^{x,v}) ds \\ &\quad + \partial_v \mu^\eta(t-s, Y_s^{x,v}, W_s^{x,v}) dB_s, \end{aligned}$$

and since μ^η is a strong solution of (21), we get precisely that for any $0 \leq s \leq s' \leq t$:

$$\theta_{s'} - \theta_s = \int_s^{s'} e^u \partial_v \mu^\eta(t-u, Y_u^{x,v}, W_u^{x,v}) dB_u.$$

In particular, $(\theta_s)_{0 \leq s \leq t}$ is a martingale, so that

$$\mu^\eta(t, x, v) = \theta_0 = \mathbb{E}[\theta_t] = e^t \mathbb{E}[f_0 * \tilde{\chi}_\eta(Y_t^{x,v}, W_t^{x,v})]. \quad (25)$$

Step 2. Proof in the case of uniform exponential tails.

But by the hypothesis on f_0 , and since χ has support in $[-1, 1]$,

$$\begin{aligned} f_0 * \tilde{\chi}_\eta(x, v) &= \int f_0(x-y, v-w) \tilde{\chi}_\eta(y, w) dy dw \leq \|f_0\|_{e,\lambda} \int_{\mathbb{R}} e^{-\lambda|v-\eta w'|} \chi(w') dw' \\ &\leq \|f_0\|_{e,\lambda} e^{-\lambda|v|} \int_{\mathbb{R}} e^{\lambda\eta|w'|} \chi(w') dw' \leq \|f_0\|_{e,\lambda} e^{-\lambda|v|+\lambda\eta}. \end{aligned} \quad (26)$$

Moreover, the definition (24) of W^η also implies that for $0 \leq s \leq t$:

$$W_s^{x,v} = e^s v - \int_0^s e^{s-u} \widetilde{K}_\eta[\mu_u^\eta] du + M_s, \quad \text{with } M_s := \sqrt{2} \int_0^s e^{s-u} dB_u,$$

Remark that M_s is in fact a centered Gaussian variable with variance $e^{2s} - 1$. Since $\|\widetilde{K}_\eta[\mu_u^\eta]\|_\infty \leq 1/2$, it leads to the following lower bound

$$|W_t^{x,v}| \geq e^t |v| - |M_t| - \frac{e^t}{2}. \quad (27)$$

Using all of this in the representation formula (25) leads to

$$\mu_t^\eta(x, v) \leq \|f_0\|_{e,\lambda} e^{t+\lambda\eta} \mathbb{E} \left[e^{-\lambda |W_t^{x,v}|} \right] \leq \|f_0\|_{e,\lambda} e^{t+\lambda(e^t/2+\eta)-\lambda e^t |v|} \mathbb{E} \left[e^{\lambda |M_t|} \right]$$

An application of (14) to M_t leads to the following bound that is uniform in η (for η small) :

$$\|\mu_t^\eta\|_{e,\lambda e^t} \leq 2 e^{t+\lambda\eta+\frac{\lambda}{2}e^t+\frac{\lambda^2}{2}} \|f_0\|_{e,\lambda}. \quad (28)$$

The conclusion follow by replacing λ by λe^{-t} in the above bound. And the estimate on $\|\rho_t^\eta\|_\infty$ is simply obtained by integration on v .

Step 3. Proof in the case of uniform polynomial tails.

The simple inequality $\langle v+w \rangle \leq \sqrt{2} \langle v \rangle \langle w \rangle$ (recall that $\langle v \rangle^2 := 1 + v^2$) implies that $\langle v-w \rangle^{-1} \leq \sqrt{2} \langle v \rangle^{-1} \langle w \rangle$, which allows to bound

$$\begin{aligned} f_0 * \tilde{\chi}_\eta(x, v) &= \int f_0(x-y, v-w) \tilde{\chi}_\eta(y, w) dy dw \leq \|f_0\|_{p,\gamma} \int_{\mathbb{R}} \langle v-\eta w' \rangle^{-\gamma} \chi(w') dw', \\ &\leq 2^{\gamma/2} \|f_0\|_{p,\gamma} \langle v \rangle^{-\gamma} \int_{\mathbb{R}} \langle \eta w' \rangle^\gamma \chi(w') dw' \leq 2^{\gamma/2} \|f_0\|_{p,\gamma} \langle \eta \rangle^\gamma \langle v \rangle^{-\gamma}. \end{aligned}$$

Plugging it into the representation formula (25), using the lower bound (27) and the simple inequality, $\langle v + \frac{1}{2}e^t \rangle^\gamma \leq 2^{\gamma/2} (e^{\gamma t} + (2v)^\gamma)$ (simply separate the cases $v \leq \frac{1}{2}e^t$ and $v \geq \frac{1}{2}e^t$),

$$\begin{aligned} \mu^\eta(t, x, v) &= e^t \mathbb{E} [f_0 * \tilde{\chi}_\eta(Y_t^{x,v}, W_t^{x,v})] \leq 2^{\gamma/2} \|f_0\|_{p,\gamma} \langle \eta \rangle^\gamma \mathbb{E} [\langle W_t^{x,v} \rangle^{-\gamma}] \\ &\leq 2^{\gamma/2} \|f_0\|_{p,\gamma} \langle \eta \rangle^\gamma \mathbb{E} \left[\left\langle e^t |v| - |M_t| - \frac{e^t}{2} \right\rangle^{-\gamma} \right] \\ &\leq 2^\gamma \|f_0\|_{p,\gamma} \langle \eta \rangle^\gamma \langle e^t v \rangle^{-\gamma} \mathbb{E} \left[\left\langle |M_t| + \frac{e^t}{2} \right\rangle^\gamma \right] \\ &\leq 2^{3\gamma/2} \|f_0\|_{p,\gamma} \langle \eta \rangle^\gamma \langle v \rangle^{-\gamma} \left(e^{\gamma t} + 2^\gamma \mathbb{E} [|M_t|^\gamma] \right). \end{aligned}$$

Since $M_t \sim \mathcal{N}(0, e^{2t}-1)$ we have $\mathbb{E} [|M_t|^\gamma] = (e^{2t}-1)^{\frac{\gamma}{2}} m_\gamma \leq e^{\gamma t} m_\gamma$, where m_γ stands for the moment of order γ of the law $\mathcal{N}(0, 1)$. This implies the claimed bound on $\|\mu_t^\eta\|_{p,\gamma}$ with $C_\gamma := 2^{3\gamma/2} (1+2^\gamma m_\gamma)$. The bound on $\|\rho_t^\eta\|_\infty$ follows from the simple bound $\langle v \rangle^{-\gamma} \leq \min(1, |v|^{-\gamma})$ and an integration in v . \square

Completeness estimates. Thanks to the propagation of the uniform estimates on the tails in velocity and the crucial Lemma 4, we are in position to show that the family $(Y_t^\eta, W^\eta)_{\eta>0}$ of solutions to (19) has the Cauchy property.

Lemma 6. For $\eta, \eta' > 0$, let $(Y_t^\eta, W_t^\eta)_{t \geq 0}$ and $(Y_t^{\eta'}, W_t^{\eta'})_{t \geq 0}$ be two (unique) solutions of the non-linear SDE (19), constructed on a given probability basis, with a common Brownian motion $(B_t)_{t \geq 0}$, and initial condition chosen as (20). If the law of the initial condition satisfies either $\|f_0\|_{e, \lambda} < \infty$ for some $\lambda > 0$ or $\|f_0\|_{e, \gamma} < \infty$ for some $\gamma > 1$, then the following stability estimate holds for any $t \geq 0$:

$$\mathbb{E} \left[\sup_{s \in [0, t]} \left(|Y_s^\eta - Y_s^{\eta'}| + |W_s^\eta - W_s^{\eta'}| \right) \right] \leq \frac{3}{2}(\eta + \eta') \exp\left((1 + 8K_{t, \eta + \eta'})t\right), \quad (29)$$

where $K_{t, \eta + \eta'}$ is the constant appearing in the r.h.s of (23) with η replaced by $\eta + \eta'$, i.e. respectively

$$K_{t, \eta + \eta'} = \frac{4}{\lambda} e^{t + \frac{\lambda}{2} + \frac{\lambda^2}{2}} e^{\lambda(\eta + \eta')} \|f_0\|_{e, \lambda} e^{-t}, \quad \text{or} \quad K_{t, \eta + \eta'} = \frac{2\gamma}{\gamma - 1} C_\gamma e^{\gamma t} \langle \eta + \eta' \rangle^\gamma \|f_0\|_{p, \gamma}.$$

Proof. Our strategy is the same as in the proof of the weak strong stability estimate in subsection 3.1. For $t \in [0, T]$ we have:

$$\begin{aligned} \sup_{s \in [0, t]} |Y_s^\eta - Y_s^{\eta'}| &\leq \int_0^t |W_s^\eta - W_s^{\eta'}| ds + |\eta - \eta'| |U| \\ \sup_{s \in [0, t]} |W_s^\eta - W_s^{\eta'}| &\leq \int_0^t \left| \tilde{K}_\eta[\mu_s^\eta](Y_s^\eta) - \tilde{K}_{\eta'}[\mu_s^{\eta'}](Y_s^{\eta'}) \right| ds + |\eta - \eta'| |V| \\ &\leq \int_0^t \left(\left| \tilde{K}_\eta[\mu_s^\eta](Y_s^\eta) - \tilde{K}[\mu_s^\eta](Y_s^\eta) \right| + \left| \tilde{K}[\mu_s^\eta](Y_s^\eta) - \tilde{K}[\mu_s^{\eta'}](Y_s^{\eta'}) \right| \right. \\ &\quad \left. + \left| \tilde{K}[\mu_s^{\eta'}](Y_s^{\eta'}) - \tilde{K}_{\eta'}[\mu_s^{\eta'}](Y_s^{\eta'}) \right| \right) ds + |\eta - \eta'|, \end{aligned}$$

since $|V|$ and $|U|$ are always bounded by 1 (we recall that χ has its support included in $[-1, 1]$). But thanks to (18), for any $y \in \mathbb{R}$:

$$\left| \tilde{K}_\eta[\mu_s^\eta](y) - \tilde{K}[\mu_s^\eta](y) \right| = \left| \int (K_\eta - K)(y - y') \rho_s^\eta(dy') \right| \leq \int \mathbb{I}_{[-\eta, \eta]}(y - y') \rho_s^\eta(dy') \leq 2\eta \|\rho_s^\eta\|_\infty.$$

This allows to bound the the first and third term, in the r.h.s. of the second inequality by $2 \int_0^t (\eta \|\rho_s^\eta\|_\infty + \eta' \|\rho_s^{\eta'}\|_\infty) ds$. And the second term is estimated in expectation with the help of Lemma 4:

$$\begin{aligned} \mathbb{E} \left[\left| \tilde{K}[\mu_s^\eta](Y_s^\eta) - \tilde{K}[\mu_s^{\eta'}](Y_s^{\eta'}) \right| \right] &\leq 8 \min(\|\rho_t^\eta\|_\infty, \|\rho_t^{\eta'}\|_\infty) \mathbb{E}[|Y_t^\eta - Y_t^{\eta'}|] \\ &\leq 4(\|\rho_t^\eta\|_\infty + \|\rho_t^{\eta'}\|_\infty) \mathbb{E}[|Y_t^\eta - Y_t^{\eta'}|]. \end{aligned}$$

Gathering all of this leads to

$$\mathbb{E} \left[\sup_{s \in [0, t]} \left(|Y_s^\eta - Y_s^{\eta'}| + |W_s^\eta - W_s^{\eta'}| \right) \right] \leq \int_0^t \alpha(s) \left(\mathbb{E}[|Y_s^\eta - Y_s^{\eta'}| + |W_s^\eta - W_s^{\eta'}|] + \frac{3}{2}(\eta + \eta') \right) ds,$$

with $\alpha(s) := 1 + 4\|\rho_s^\eta\|_\infty + 4\|\rho_s^{\eta'}\|_\infty$. A simple application of the Gronwall's lemma to $\mathbb{E}[\sup_{s \in [0, t]} (|Y_s^\eta - Y_s^{\eta'}| + |W_s^\eta - W_s^{\eta'}|) + 3(\eta + \eta')/2]$ leads to

$$\mathbb{E} \left[\sup_{s \in [0, t]} \left(|Y_s^\eta - Y_s^{\eta'}| + |W_s^\eta - W_s^{\eta'}| \right) \right] \leq \frac{3}{2}(\eta + \eta') \exp\left(\int_0^t \alpha(s) ds\right).$$

The conclusion follows from (23), which implies that $\|\rho_s^\eta\|_\infty \leq K_{s, \eta} \leq K_{t, \eta + \eta'}$ for any $s \in [0, t]$ (and a similar inequality with η' replacing η). \square

The Cauchy property in the space of path. We now consider the space \mathcal{A} of measurable applications (or random variables) from $\Omega = C(\mathbb{R}^+; \mathbb{R}^2)$ (with the Wiener measure) into itself. We endow it with the topology of uniform convergence on compact (in time) subsets. An associated distance to that topology is for instance

$$\forall Y, Z \in \mathcal{A}, \quad d(Y, Z) = \sum_{n \in \mathbb{N}^*} \frac{1}{2^n} \mathbb{E} \left(1 \wedge \sup_{t \in [0, 2^n]} (|Y_t^1 - Z_t^1| + |Y_t^2 - Z_t^2|) \right), \quad (30)$$

for which \mathcal{A} is complete (the symbol \wedge stand for the minimum). Let us consider the “sequence” $(Y^\eta, W^\eta)_{\eta > 0}$ (the correct denomination is “net”). By Lemma 6 it is a Cauchy “sequence” (or net) in (\mathcal{A}, d) , and then it converges towards a certain $(Y_t, W_t)_{t \geq 0}$ in (\mathcal{A}, d) .

At a fixed time t , this implies the convergence in probability and then in law of (Y_t^η, W_t^η) toward (Y_t, W_t) : *i.e.* the time marginals μ_t^η weakly converge (as measures) towards f_t . Using a standard argument, we can pass in the limit in the uniform bound (22) obtained in Lemma 5. So that for any time t , the density of the law f_t of (Y_t, W_t) satisfies one of the bound of (6).

Identification of the limit. In order to prove that $(Y_t, W_t)_{t \geq 0}$ is a solution to (3) we have to show that for any $t \geq 0$:

$$\mathbb{E} \left[\sup_{s \in [0, t]} \left| Y_s - Y_0 - \int_0^s W_u du \right| \right] = 0, \quad \mathbb{E} \left[\sup_{s \in [0, t]} \left| W_s - W_0 + \int_0^s (W_u - \tilde{K}[\mu_u](Y_u)) du - B_s \right| \right] = 0. \quad (31)$$

But this is something we know for the approximated process $(Y_t^\eta, W_t^\eta)_{t \geq 0}$. Precisely, by Definition of strong solution to (19):

$$\begin{aligned} \mathbb{E} \left[\sup_{s \in [0, t]} \left| Y_s^\eta - Y_0 - \eta U - \int_0^s W_u^\eta du \right| \right] &= 0, \\ \mathbb{E} \left[\sup_{s \in [0, t]} \left| W_s^\eta - W_0 - \eta V + \int_0^s (W_u^\eta - \tilde{K}_\eta[\mu_u^\eta](Y_u^\eta)) du - B_s \right| \right] &= 0. \end{aligned}$$

But the convergence in (\mathcal{A}, d) allows to pass to the limit in the first equality above , and we obtain the first equality in (31). In order to pass to the limit in the second inequality, and get the second part of (31), the main difficulty is to handle the non-linear term. Precisely, we will show in the rest of the proof that

$$\sup_{s \in [0, t]} \mathbb{E} \left(\left| \tilde{K}_\eta[\mu_s^\eta](Y_s^\eta) - \tilde{K}[\mu_s](Y_s) \right| \right) \xrightarrow{\eta \rightarrow 0} 0.$$

This will conclude the proof.

For this, we introduce a independent copy $(\bar{Y}_t^\eta, \bar{Y}_t)$ of the couple (Y_t^η, Y_t) . We then rewrite the force with the help of that independent copy and estimate

$$\begin{aligned} \mathbb{E} \left[\left| \tilde{K}_\eta[\mu_s^\eta](Y_s^\eta) - \tilde{K}[\mu_s](Y_s) \right| \right] &= \mathbb{E} \left[\left| K_\eta(Y_s^\eta - \bar{Y}_s^\eta) - K(Y_s - \bar{Y}_s) \right| \right] \\ &\leq \mathbb{E} \left[\left| K_\eta - K \right| (Y_s^\eta - \bar{Y}_s^\eta) \right] + \mathbb{E} \left[\left| K(Y_s^\eta - \bar{Y}_s^\eta) - K(Y_s - \bar{Y}_s) \right| \right] \end{aligned}$$

The first term in the r.h.s is bounded thanks to (18):

$$\begin{aligned} \mathbb{E} \left[\left| K_\eta - K \right| (Y_s^\eta - \bar{Y}_s^\eta) \right] &\leq \mathbb{E} \left[\mathbb{I}_{[-\eta, \eta]}(Y_s^\eta - \bar{Y}_s^\eta) \right] \leq \mathbb{E} \left[\mathbb{E} \left[\mathbb{I}_{[-\eta, \eta]}(Y_s^\eta - \bar{Y}_s^\eta) | Y_s^\eta \right] \right] \\ &\leq \mathbb{E} [2\eta \|\rho_s^\eta\|_\infty] = 2\eta \|\rho_s^\eta\|_\infty \end{aligned}$$

The second term in the r.h.s. is bounded thanks to Lemma 4 and we get

$$\mathbb{E} \left[\left| \tilde{K}_\eta[\mu_s^\eta](Y_s^\eta) - \tilde{K}[\mu_s](Y_s) \right| \right] \leq 2\eta \|\rho_s^\eta\|_\infty + 8 \|\rho_s\|_\infty \mathbb{E}[|Y_s^\eta - Y_s|].$$

Since $\|\rho_s^\eta\|_\infty$ and $\|\rho_s\|_\infty$ are bounded uniformly in time on $[0, t]$ and for $\eta \in (0, 1)$, and thanks to the convergence of (Y^η, W^η) towards (Y, W) in (\mathcal{A}, d) , it is simple to conclude that the requested term goes to zero, as η goes to zero.

3.4 Proof of Corollary 1

The existence part is a simple consequence of the existence of the process solution to the non linear SDE (3), for a given initial condition with a polynomial or exponential decays of the velocity tails (use Ito's rule). So we only prove the weak-strong stability estimate.

Let $(g_t)_{t \geq 0}$ be weak solution to (8) starting from $g_0 \in \mathcal{P}_1(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)$ and $(f_t)_{t \geq 0}$ be another weak solution to (8) starting from $f_0 \in \mathcal{P}_1(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)$ satisfying for any $t > 0$

$$\sup_{s \leq t} \|\rho_s\|_\infty < \infty, \quad \text{where} \quad \rho_s(x) = \rho_s^f := \int_{\mathbb{R}} f_s(x, dv).$$

By [7, Theorem 2.6] or [9, Proposition B.1], there exists a stochastic basis $(\Omega, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F})$ and a Brownian motion $(B_t)_{t \geq 0}$ on this basis and a process $(X_t, V_t)_{t \geq 0}$ solution to (3), which has exactly the time marginal g_t at any time $t \geq 0$.

Next, remark that the force field \tilde{K} created by f , precisely $\tilde{K}_s(x) = \int_{\mathbb{R}} K(x - y) \rho_s^f(dy)$, is Lipschitz in position. In fact,

$$|\tilde{K}_s(x) - \tilde{K}_s(y)| \leq \int_{\mathbb{R}} |K(x - z) - K(y - z)| \rho_s(dz) \leq \int_{\mathbb{R}} \mathbb{I}_{[x, y]}(z) \rho_s(dz) \leq \|\rho_s\|_\infty |x - y|.$$

Extending the probability space, we may choose a r.v. $(Y_0, W_0) \sim f_0$ such that

$$W_1(f_0, g_0) = \mathbb{E} \left[|Y_0 - X_0| + |W_0 - V_0| \right],$$

and since \tilde{K} is regular, standard results allow us to build on that probability space a stochastic process $(Y_t, W_t)_{t \geq 0}$ solution to

$$Y_t = Y_0 + \int_0^t W_s ds, \quad W_t = W_0 + \int_0^t \tilde{K}_s(Y_s) ds - \int_0^t W_s ds + \sqrt{2} B_t.$$

Note that the family of time marginals $(h_t)_{t \geq 0}$ of (Y_t, W_t) is a solution to the following linear Vlasov-Fokker-Planck equation

$$\partial_t h_t + v \partial_x h_t + \tilde{K}_t(x) \partial_v h_t = \partial_v (\partial_v h_t + v h_t), \quad \text{where} \quad \tilde{K}_s(x) = K * \rho_s^f(x), \quad (32)$$

with initial condition $h_{t=0} = f_0$. Of course, that equation is also satisfied by f by assumption. But since \tilde{K} is globally Lipschitz in the space variable, uniformly in the time variable, uniqueness holds for equation (32) in the class $L_t^\infty(L^1)$ by standard results. So, we have $h_t = f_t$ for all time $t \geq 0$ and $(Y_t, W_t)_{t \geq 0}$ is actually a solution to (3) defined on the same probability space as $(X_t, V_t)_{t \geq 0}$. Then applying the point (ii) of Theorem 2 to those processes and recalling that

$$W_1(f_t, g_t) \leq \mathbb{E} \left[|Y_t - X_t| + |W_t - V_t| \right],$$

we get the expected estimate.

4 Proof of Theorem 3

Before going to the proof, we will prove a useful lemma.

Lemma 7. *Let (X_1, \dots, X_N) be N i.i.d random variables of law $\rho \in \mathcal{P}(\mathbb{R}^d)$, and $\rho_N = \frac{1}{N} \sum_i \delta_{X_i}$ be the associated empirical measure. Then, for all $a \in \mathbb{R}^d$ we have:*

$$\mathbb{E} \left[\sup_{u \in \mathbb{R}_+} \left| \int_{\mathbb{R}^d} \mathbb{I}_{|a-y| \leq u} (\rho_N - \rho)(dy) \right| \right] \leq \frac{3}{\sqrt{N}}$$

Proof. In fact, we choose a sequence $(X_n)_{n \in \mathbb{N}}$ of i.i.d random variables of law $\rho \in \mathcal{P}(\mathbb{R}^d)$ and L some Poisson random variable of parameter N independent of the $(X_n)_{n \in \mathbb{N}}$. We define the two point process by

$$M_N = \sum_{i=1}^L \delta_{Y_i},$$

M_N is in fact a Poisson Random Measure (PRM) with intensity measure $N\rho$. Remark that $\|M_N - N\rho_N\|_{TV} = |L - N|$. For all $a \in \mathbb{R}^d$, we have

$$\begin{aligned} \left| \int \mathbb{I}_{|a-y| \leq u} (\rho_N - \rho)(dy) \right| &\leq \left| \int \mathbb{I}_{|a-y| \leq u} \left(\frac{1}{N} M_N - \rho \right)(dy) \right| + \left| \int \mathbb{I}_{|a-y| \leq u} \left(\frac{1}{N} M_N - \rho_N \right)(dy) \right| \\ &\leq \frac{1}{N} \left| \int \mathbb{I}_{|a-y| \leq u} (M_N - N\rho)(dy) \right| + \frac{1}{N} \|M_N - N\rho_N\|_{TV} \\ \sup_{u \in \mathbb{R}_+} \left| \int \mathbb{I}_{|a-y| \leq u} (\rho_N - \rho)(dy) \right| &\leq \frac{1}{N} \sup_{u \in \mathbb{R}_+} |\mathcal{M}_u^{N,a}| + \frac{|L - N|}{N}, \end{aligned} \quad (33)$$

$$\text{where } \mathcal{M}_u^{N,a} = \int \mathbb{I}_{|a-y| \leq u} (M_N - N\rho)(dy).$$

Since M_N is a PRM, $(\mathcal{M}_u^{N,a})_{u \geq 0}$ is a martingale with respect to the filtration $(\mathcal{F}_u^a)_{u \geq 0} = (\sigma(M_N \mathbb{I}_{B(a,u)}))_{u \geq 0}$, where $B(a, u)$ the ball of center a and radius u in \mathbb{R}^d . So using Doob's inequality and the fact that $\mathcal{M}_\infty^{N,a} = L - N$ we find

$$\mathbb{E} \left[\sup_{u \geq 0} |\mathcal{M}_u^{N,a}| \right] \leq \left(\mathbb{E} \left[\sup_{u \geq 0} |\mathcal{M}_u^{N,a}|^2 \right] \right)^{1/2} \leq 2 \left(\mathbb{E} [|\mathcal{M}_\infty^{N,a}|^2] \right)^{1/2} \leq 2 \mathbb{V}[L] = 2\sqrt{N}.$$

Taking now the expectation in (33), using the above bound and $\mathbb{E}|L - N| \leq \sqrt{\mathbb{V}(L)} = \sqrt{N}$, we conclude the proof. \square

We are now in position to prove Theorem 3.

Step 1. Coupling estimates.

We begin by a calculation valid for any fixed realization of these processes (*i.e.* given any initial conditions and Brownian paths): for all $i = 1, \dots, N$ we have:

$$\begin{aligned} \sup_{s \in [0, t]} |X_{i,s}^N - Y_{i,s}^N| &\leq |X_{i,0}^N - Y_{i,0}^N| + \int_0^t |V_{i,s}^N - W_{i,s}^N| ds \\ \sup_{s \in [0, t]} |V_{i,s}^N - W_{i,s}^N| &\leq |V_{i,0}^N - W_{i,0}^N| + \int_0^t \left| \frac{1}{N} \sum_{j=1}^N K(X_{i,s}^N - X_{j,s}^N) - \left(\int_{\mathbb{R} \times \mathbb{R}} K(Y_{i,s}^N - x) \mu_t(dx, dv) \right) \right| \\ &\leq |V_{i,0}^N - W_{i,0}^N| + \int_0^t \left| \frac{1}{N} \sum_{j=1}^N K(X_{i,s}^N - X_{j,s}^N) - K(Y_{i,s}^N - Y_{j,s}^N) \right| ds + \frac{t}{N-1} + \int_0^t \Lambda_{i,s}^N ds \\ \text{with } \Lambda_{i,s}^N &:= \left| \frac{1}{N-1} \sum_{j \neq i}^N K(Y_{i,s}^N - Y_{j,s}^N) - \int_{\mathbb{R} \times \mathbb{R}} K(Y_{i,s}^N - x) \mu_s(dx, dv) \right| \end{aligned}$$

We sum these inequalities over $i = 1, \dots, N$, divide by $N \geq 2$ and then get:

$$\begin{aligned}
\frac{1}{N} \sum_{i=1}^N \sup_{s \in [0, t]} |X_{i,s}^N - Y_{i,s}^N| &\leq \frac{1}{N} \sum_{i=1}^N |X_{i,0}^N - Y_{i,0}^N| + \int_0^t \frac{1}{N} \sum_{i=1}^N |V_{i,s}^N - W_{i,s}^N| ds \\
\frac{1}{N} \sum_{i=1}^N \sup_{s \in [0, t]} |V_{i,s}^N - W_{i,s}^N| &\leq \frac{1}{N} \sum_{i=1}^N |V_{i,0}^N - W_{i,0}^N| + \frac{2t}{N} + \int_0^t \Lambda_s^N ds \\
&\quad + \frac{1}{N^2} \sum_{i \neq j}^N \int_0^t |K(X_{i,s}^N - X_{j,s}^N) - K(Y_{i,s}^N - Y_{j,s}^N)| ds
\end{aligned} \tag{34}$$

where $\Lambda_s^N := \frac{1}{N} \sum_{i=1}^N \Lambda_{i,s}^N ds$.

Using equality (11) we find with the notation $\hat{\rho}_s^{i,N} = \frac{1}{N-1} \sum_{j \neq i} \delta_{Y_{j,s}^N}$,

$$\begin{aligned}
&\frac{1}{N^2} \sum_{i \neq j}^N |K(X_{i,s}^N - X_{j,s}^N) - K(Y_{i,s}^N - Y_{j,s}^N)| \\
&\leq \frac{2}{N^2} \sum_{i \neq j}^N \mathbb{I}_{|Y_{i,s}^N - Y_{j,s}^N| \leq 2|X_{i,s}^N - Y_{i,s}^N|} = \frac{2(N-1)}{N^2} \sum_{i=1}^N \int_{\mathbb{R}} \mathbb{I}_{|Y_{i,s}^N - y| \leq 2|X_{i,s}^N - Y_{i,s}^N|} \hat{\rho}_s^{i,N}(dy) \\
&\leq \frac{2}{N} \sum_{i=1}^N \int_{\mathbb{R}} \mathbb{I}_{|Y_{i,s}^N - y| \leq 2|X_{i,s}^N - Y_{i,s}^N|} \rho_s(dy) + \frac{2}{N} \sum_{i=1}^N \left| \int_{\mathbb{R}} \mathbb{I}_{|Y_{i,s}^N - y| \leq 2|X_{i,s}^N - Y_{i,s}^N|} (\hat{\rho}_s^{i,N} - \rho_s)(dy) \right| \\
&\leq \frac{8}{N} \|\rho_s\|_{\infty} \sum_{i=1}^N |X_{i,s}^N - Y_{i,s}^N| + \Gamma_s^N
\end{aligned}$$

where

$$\Gamma_s^N := \frac{2}{N} \sum_{i=1}^N \Gamma_{i,s}^N, \quad \Gamma_{i,s}^N := \sup_{u \in \mathbb{R}_+} \left| \int_{\mathbb{R}} \mathbb{I}_{|Y_{i,s}^N - y| \leq u} (\hat{\rho}_s^{i,N} - \rho_s)(dy) \right|. \tag{35}$$

Then, $\beta(t) := \frac{1}{N} \sum_{i=1}^N (\sup_{s \in [0, t]} |X_{i,s}^N - Y_{i,s}^N| + \sup_{s \in [0, t]} |V_{i,s}^N - W_{i,s}^N|)$ satisfies the integral inequality

$$\beta(t) \leq \beta(0) + \int_0^t \left((1 + 8\|\rho_s\|_{\infty}) \beta(s) + \Lambda_s^N + 2\Gamma_s^N + \frac{2}{N} \right) ds.$$

An application of Grönwall's Lemma leads to

$$\beta(t) \leq e^{(t+8 \int_0^t \|\rho_s\|_{\infty} ds)} \left(\beta(0) + \int_0^t e^{-s} \left(\Lambda_s^N + 2\Gamma_s^N + \frac{2}{N} \right) ds \right).$$

Taking the expectation and using the symmetry of the laws of the of $(X_{i,t}^N, V_{i,t}^N)_{i=1, \dots, N}$ and $(Y_{i,t}^N, W_{i,t}^N)_{i=1, \dots, N}$ we find

$$\begin{aligned}
\mathbb{E} \left[\sup_{s \in [0, t]} |X_{1,s}^N - Y_{1,s}^N| + |V_{1,s}^N - W_{1,s}^N| \right] &\leq e^{(t+8 \int_0^t \|\rho_s\|_{\infty} ds)} \\
&\quad \left(\mathbb{E} \left[|X_{1,0}^N - Y_{1,0}^N| + |V_{1,0}^N - W_{1,0}^N| \right] + \int_0^t e^{-s} \left(\mathbb{E} [\Lambda_{1,s}^N + 2\Gamma_{1,s}^N] + \frac{2}{N} \right) ds \right). \tag{36}
\end{aligned}$$

We will bound the expectation of stochastic terms appearing in the r.h.s. with the help of Lemma 7.

Step 2. Conclusion of the proof.

We recall that $\hat{\rho}_t^{i,N} = \frac{1}{N-1} \sum_{j \neq i} \delta_{Y_{j,t}^N}$ is the empirical measure associated to the $(Y_{i,t}^N)_{2 \leq i \leq N}$, and is then independent of $Y_{1,t}^N$. By the definition (35) of Γ and Lemma 7, we have

$$\begin{aligned} \mathbb{E} [\Gamma_{1,t}^N] &= \mathbb{E} \left[\sup_{u \geq 0} \left| \int \mathbb{I}_{|Y_{1,t}^N - y| \leq u} (\hat{\rho}_t^{1,N} - \rho_t)(dy) \right| \right], \\ &= \mathbb{E} \left[\mathbb{E} \left[\sup_{u \geq 0} \left| \int \mathbb{I}_{|Y_{1,t}^N - y| \leq u} (\hat{\rho}_t^{1,N} - \rho_t)(dy) \right| \middle| Y_{1,t}^N \right] \right] \leq \frac{3}{\sqrt{N}}, \end{aligned}$$

Moreover, using again the fact that the $(Y_{i,t}^N)_{1 \leq i \leq N}$ are i.i.d and that $\|K\|_\infty \leq 1/2$, we find for $N \geq 2$

$$\begin{aligned} \mathbb{E} [\Lambda_{1,t}^N] &= \mathbb{E} \left[\mathbb{E} \left[\left| \frac{1}{N-1} \sum_{j=2}^N K(Y_{1,t}^N - Y_{j,t}^N) - \int_{\mathbb{R}^2} K(Y_{1,t}^N - x) \mu_s(dx, dv) \right| \middle| Y_{1,t}^N \right] \right], \\ &\leq \mathbb{E} \left[\left(\mathbb{V} \left[\frac{1}{N-1} \sum_{j=2}^N K(Y_{1,t}^N - Y_{j,t}^N) \middle| Y_{1,t}^N \right] \right)^{1/2} \right] \leq \frac{1}{2\sqrt{N-1}} \leq \frac{1}{\sqrt{2N}}. \end{aligned}$$

Then applying these results to equation (36) leads to

$$\mathbb{E} \left[\sup_{s \in [0,t]} |X_{1,s}^N - Y_{1,s}^N| + |V_{1,s}^N - W_{1,s}^N| \right] \leq e^{\left(t + 8 \int_0^t \|\rho_s\|_\infty ds \right)} \left(\frac{1}{\sqrt{N}} (1 + 2 \times 3) + \frac{2}{N} \right),$$

which leads to the claimed bound.

5 Proof of Proposition 2

We distinguish here the deviation upper bounds for $\sup_{t \in [0,T]} \|\mu_{Y,t}^N\|_{\infty,\varepsilon}$ because such a result have an interest by itself. Moreover, we will see it is used in the proof of the concentration inequalities. In the present proof, we apply the following strategy: we first find prove concentration inequalities for the quantity $\|\mu_{Y,t}^N\|_{\infty,\varepsilon}$ at a fixed time t , then for the variation of this quantity on small time intervals, and finally conclude by mixing both estimates in an optimal way. That proof could be extended to dimension larger than 1, but we prefer to restrict here to the case of interest.

5.1 A uniform deviation upper bound for Binomial variables

We begin with a general result about deviation upper bounds for Binomial variable.

Lemma 8. *Let $p \in [0, 1]$. If X is binomial variable of parameter (N, p) , then for any $\alpha > 0$:*

$$\mathbb{P}(|X - Np| \geq N\alpha) \leq 2e^{-2\alpha^2 N}.$$

Proof. We may write $X = \sum_{i=1}^N X_i$, with a family $(X_i)_{i \leq N}$ of N i.i.d. Bernoulli variables of parameter p . For all $\lambda > 0$ we have:

$$\mathbb{E}[e^{\lambda X}] = \prod_{i=1}^N \mathbb{E}[e^{\lambda X_i}] = \left(\mathbb{E}[e^{\lambda X_1}] \right)^N = \left(1 - p + pe^\lambda \right)^N.$$

Using Markov's inequality we get:

$$\mathbb{P}(X \geq N(p + \alpha)) \leq \left(\frac{1 - p + pe^\lambda}{e^{\lambda(p + \alpha)}} \right)^N = \left((1 - p)e^{-\lambda(p + \alpha)} + pe^{\lambda(1 - (p + \alpha))} \right)^N,$$

The optimal λ turns out to be $\ln \left(\frac{(p+\alpha)(1-p)}{(1-(p+\alpha))p} \right)$, and we get after some calculations:

$$\mathbb{P}(X \geq N(p+\alpha)) \leq e^{-Ng_p(\alpha)},$$

where

$$g_p(\alpha) := H(\mathcal{B}(p+\alpha) | \mathcal{B}(p)) = (1-(p+\alpha)) \ln \left(\frac{1-(p+\alpha)}{1-p} \right) + (p+\alpha) \ln \left(\frac{p+\alpha}{p} \right),$$

is the relative entropy with respect to $\mathcal{B}(p)$. By the properties of relative entropy, we have $g_p(0) = 0$ and $g_p(\alpha) > 0$ if $\alpha \neq 0$. So $g'_p(0) = 0$. Moreover a straight calculation gives for all α such that $p+\alpha < 1$:

$$g''_p(\alpha) = \frac{1}{p+\alpha} + \frac{1}{1-(p+\alpha)} \geq 4,$$

since for all $x \in (0, 1)$, $x(1-x) \leq \frac{1}{4}$. Using Taylor's formula with integral rest at order 2 we have:

$$g_p(\alpha) = g_p(0) + g'_p(0)\alpha + \int_0^\alpha g''_p(u)(\alpha-u)du \geq 2\alpha^2.$$

So combining all these estimates we get for $\alpha + p < 1$:

$$\mathbb{P}(X \geq N(p+\alpha)) \leq e^{-Ng_p(\alpha)} \leq e^{-2N\alpha^2}.$$

It is still valid for $\alpha + p = 1$: just pass to the limit, and there is nothing to prove for $p+\alpha > 1$, since X cannot be larger than N . Moreover, since $\mathbb{P}(X \leq N(p-\alpha)) = \mathbb{P}(N-X \geq N((1-p)+\alpha))$, an application of the above bound to the $\mathcal{B}(N, 1-p)$ -Binomial variable $N-X$ leads to

$$\mathbb{P}(X \leq N(p-\alpha)) \leq e^{-2N\alpha^2},$$

and this concludes the proof. \square

5.2 Concentration inequalities at fixed time

Thanks to lemma 8 we are able to give some concentration inequalities for the empirical measure $\rho^N = \frac{1}{N} \sum_{i=1}^N \delta_{Y_i}$, for i.i.d.r.v. $(Y_i)_{i \in N}$.

Lemma 9. *Let $\alpha, \varepsilon > 0$. Assume that (Y_1, \dots, Y_N) are N independent random variables, all with law $\rho \in L^\infty$ (we identify the law and its density). Assume also that ρ has an exponential moment of order $\lambda > 0$: $\mathcal{M}_\lambda(\rho) := \int e^{\lambda|y|} \rho(dy) < +\infty$. We denote by $\rho^N := \frac{1}{N} \sum_i \delta_{Y_i}$ the associated empirical measure. Then for any $\varepsilon, \alpha > 0$,*

$$\mathbb{P}(\|\rho^N\|_{\infty, \varepsilon} \geq \|\rho\|_\infty + \alpha) \leq \left(\frac{4\|\rho\|_\infty N(\varepsilon\alpha)}{\lambda} + 2 + N\mathcal{M}_\lambda(\rho) \right) e^{-2N(\varepsilon\alpha)^2}$$

Proof. Step 1. A first bound valid on compact subset. For any $0 < \delta < \varepsilon$ we set $k = \lfloor \frac{R}{2\delta} \rfloor + 1$. It is clear that for all $x \in [-R, R]$, there exists $\ell \in \{-k, \dots, k\}$ such that $B(x, \varepsilon) \subset B(2\ell\delta, \varepsilon + \delta)$. It implies

$$\begin{aligned} \mathbb{P}\left(\sup_{x \in [-R, R]} \frac{\rho^N[B(x, \varepsilon)]}{2\varepsilon} \geq \|\rho\|_\infty + \alpha\right) &\leq \mathbb{P}\left(\sup_{\ell = -k, \dots, k} \rho^N[B(2\ell\delta, \varepsilon + \delta)] \geq 2\varepsilon(\|\rho\|_\infty + \alpha)\right) \\ &\leq \sum_{\ell = -k}^k \mathbb{P}\left(\rho^N[B(2\ell\delta, \varepsilon + \delta)] \geq 2(\varepsilon + \delta)\|\rho\|_\infty + 2(\alpha\varepsilon - \delta\|\rho\|_\infty)\right). \end{aligned}$$

Since for any ℓ , $N\rho^N(B(2\ell\delta, \varepsilon+\delta))$ is a Binomial variable of parameter N and $p_\ell = \int_{B(2\ell\delta, \varepsilon+\delta)} \rho_t(dx) \leq 2(\varepsilon+\delta)\|\rho_t\|_\infty$, we may apply Lemma 8 and bound each term in the r.h.s. by $\exp(-8N(\alpha\varepsilon-\delta\|\rho\|_\infty)^2)$. By the definition of k , and provided that $\varepsilon\alpha \geq \delta\|\rho\|_\infty$ it leads to

$$\mathbb{P}\left(\sup_{x \in [-R, R]} \frac{\rho^N[B(x, \varepsilon)]}{2\varepsilon} \geq \|\rho\|_\infty + \alpha\right) \leq \left(\frac{R}{\delta} + 2\right)e^{-8N(\alpha\varepsilon-\delta\|\rho\|_\infty)^2}.$$

We now have to choose δ in order to minimize the right hand side. The particular choice $\delta\|\rho\|_\infty = (\varepsilon\alpha)/2$ satisfies the previous restriction and already provides an interesting bound:

$$\mathbb{P}\left(\sup_{x \in [-R, R]} \frac{\rho^N[B(x, \varepsilon)]}{2\varepsilon} \geq \|\rho\|_\infty + \alpha\right) \leq \left(\frac{2R\|\rho\|_\infty}{\varepsilon\alpha} + 2\right)e^{-2N(\varepsilon\alpha)^2}. \quad (37)$$

Step 2. Extension to the whole space. It is clear that

$$\begin{aligned} \mathbb{P}(\|\rho^N\|_{\infty, \varepsilon} \geq \|\rho\|_\infty + \alpha) &\leq \mathbb{P}\left(\sup_{x \in [-R, R]} \frac{\rho^N[B(x, \varepsilon)]}{2\varepsilon} \geq \|\rho\|_\infty + \alpha\right) + \mathbb{P}(\exists i \leq N, |Y_i| > R) \\ &\leq \left(\frac{2R\|\rho\|_\infty}{\varepsilon\alpha} + 2\right)e^{-2N(\varepsilon\alpha)^2} + N\mathcal{M}_\lambda(\rho)e^{-\lambda R}, \end{aligned}$$

where we used (37) to bound the first term in the r.h.s, and a simple application of Chebyshev's inequality to bound the second term: precisely that $\mathbb{P}(\exists i \leq N, |Y_i| > R) \leq N\mathbb{P}(|Y_1| > R) \leq N\mathcal{M}_\lambda(\rho)e^{-\lambda R}$. We choose $R = \frac{2}{\lambda}N(\alpha\varepsilon)^2$, and finally get:

$$\mathbb{P}(\|\rho^N\|_{\infty, \varepsilon} \geq \|\rho\|_\infty + \alpha) \leq \left(\frac{4\|\rho\|_\infty N(\varepsilon\alpha)}{\lambda} + 2\right)e^{-2N(\varepsilon\alpha)^2} + N\mathcal{M}_\lambda(\rho)e^{-2N(\varepsilon\alpha)^2},$$

and the result is proved. \square

5.3 Conclusion of the proof of Proposition 2

We begin with the following corollary of Proposition 3, which will be useful in the sequel

Corollary 10. *Let be $(Y_t, W_t)_{t \in [0, T]}$ be a solution of 3 for some initial condition (Y_0, W_0) of law f_0 , having an exponential moment of order $\lambda > 0$. We define $c_\lambda := \frac{5}{2} + \frac{1}{\lambda} \ln(\mathbb{E}[e^{\lambda|W_0|}])$ and denote by f_t the law of (Y_t, W_t) . For $0 \leq s < t \leq s + \min(\frac{1}{16}, \lambda^{-2})$, and $\beta > 0$, it holds:*

$$\mathbb{P}\left(\sup_{s \leq u \leq t} |Y_u - Y_s| \geq (t-s)(c_\lambda + \beta)\right) \leq e^{-\frac{\beta}{2} \min(\beta, \lambda)}.$$

And if the (Y_i^N, W_i^N) are N independent copies of the previous process, with the same notation

$$\mathbb{P}\left(\frac{1}{N} \sum_{i=1}^N \sup_{s \leq u \leq t} |Y_{i,u}^N - Y_{i,s}^N| \geq (t-s)(c_\lambda + \beta)\right) \leq e^{-\frac{\beta}{2} \min(\beta, \lambda)N}.$$

Proof. Using point (iii) of the Proposition 3, and Chebyshev's inequality, and the definition of c_λ we get for any $0 < \lambda' \leq \lambda$:

$$\begin{aligned} \mathbb{P}\left(\sup_{s \leq u \leq t} |Y_u - Y_s| \geq (t-s)(c_\lambda + \beta)\right) &\leq e^{-\lambda'(c_\lambda + \beta)} \mathbb{E}\left[e^{\lambda'(t-s)^{-1} \sup_{s \leq u \leq t} |Y_u - Y_s|}\right] \\ &\leq e^{-\lambda'(c_\lambda + \beta)} e^{\frac{1}{2}\lambda'(5+\lambda')} \mathbb{E}\left[e^{\lambda'|W_0|}\right] \\ &\leq e^{-\lambda'\beta + \frac{1}{2}(\lambda')^2} \mathbb{E}\left[e^{\lambda|W_0|}\right]^{-\frac{\lambda'}{\lambda}} \mathbb{E}\left[e^{\lambda'|W_0|}\right] \leq e^{-\lambda'\beta + \frac{1}{2}(\lambda')^2}. \end{aligned}$$

Optimization in λ' leads to the particular choice $\lambda' = \beta$, when $\beta \leq \lambda$, and to the choice $\lambda' = \lambda$ otherwise. If we use $\beta - \frac{\lambda}{2} \geq \frac{\beta}{2}$ in the later case, we obtain the expected bound.

The proof of the second bound involving N independent copies follows the same lines: by independence

$$\mathbb{P}\left(\frac{1}{N} \sum_{i=1}^N \sup_{s \leq u \leq t} |Y_{i,u}^N - Y_{i,s}^N| \geq (t-s)(c_\lambda + \beta)\right) \leq \left(e^{-\lambda'(c_\lambda + \beta)} \mathbb{E}\left[e^{\lambda'(t-s)^{-1} \sup_{s \leq u \leq t} |Y_u - Y_s|}\right]\right)^N,$$

and the conclusion follows with the same optimization on λ' . \square

We are now in position to prove Proposition 2. We fix $\gamma > 0$, define $\alpha := \frac{\gamma}{2}$, and recall from Corollary 10 the notation $c_\lambda = \frac{5}{2} + \frac{1}{\lambda} \ln\left(\mathbb{E}[e^{\lambda|W_0|}]\right)$ together with $\kappa_t := \sup_{0 \leq s \leq t} \|\rho_s\|_\infty$.

$$\beta := \sqrt{N}\varepsilon\gamma \max\left(1, \frac{\sqrt{N}\varepsilon\gamma}{\lambda}\right), \quad \Delta t := \frac{\varepsilon\alpha}{(\kappa_t + \alpha)(c_\lambda + \beta)}.$$

Remark that β satisfies $\frac{\beta}{2} \min(\beta, \lambda) = 2N(\varepsilon\alpha)^2$ and that Δt is always smaller than $\min(\frac{1}{16}, \lambda^{-2})$ by the assumptions in Proposition 2, so that we may apply Corollary 10. We then choose $K = \lfloor \frac{t}{\Delta t} \rfloor + 1$, define $t_k = k\Delta t$ for all $0 \leq k \leq K$ (remark that $t_K \geq t$), and the two following events Ω^1 and Ω^2 as:

$$\begin{aligned} \Omega^1 &= \left\{ \exists 0 \leq k \leq K-1, \quad \|\rho_{t_k}^N\|_{\infty, \varepsilon + \Delta t(c_\lambda + \beta)} > \kappa_t + \alpha \right\}, \\ \Omega^2 &= \left\{ \exists 0 \leq k \leq K-1, \exists i \leq N, \quad \sup_{s \in [t_k, t_{k+1}]} |Y_{i,s}^N - Y_{i,t_k}^N| > \Delta t(c_\lambda + \beta) \right\}. \end{aligned}$$

If the events Ω_1^c and Ω_2^c are realized, then for any $0 \leq s \leq t$, we choose k such that $s \in [t_k, t_{k+1})$ and get for any $x \in \mathbb{R}$

$$\begin{aligned} \rho_s^N(B(x, \varepsilon)) &\leq \rho_{t_k}^N(B(x, \varepsilon + \Delta t(c_\lambda + \beta))) \leq 2(\kappa_t + \alpha)(\varepsilon + \Delta t(c_\lambda + \beta)) \\ &= 2\varepsilon(\kappa_t + \alpha) \left(1 + \frac{\alpha}{(\kappa_t + \alpha)}\right) = 2\varepsilon(\kappa_t + \gamma). \end{aligned}$$

The last equalities follow from the definition of α and Δt . It means that if Ω_1^c and Ω_2^c are realized, then

$$\sup_{s \in [0, t]} \|\rho_s^N\|_{\infty, \varepsilon} \leq \kappa_t + \gamma.$$

Next, we can bound $\mathbb{P}(\Omega_1)$ and $\mathbb{P}(\Omega_2)$ with the help of Lemma 9, Corollary 10, and point (ii) of Proposition 3 (on the control of the exponential moments of ρ_t): it leads precisely to

$$\begin{aligned} \mathbb{P}\left(\sup_{s \in [0, t]} \|\rho_s^N\|_{\infty, \varepsilon} \geq \kappa_t + \gamma\right) &\leq \mathbb{P}(\Omega_1) + \mathbb{P}(\Omega_2) \\ &\leq \sum_{k=0}^K \left(\frac{4\|\rho_{t_k}\|_\infty N(\varepsilon\alpha)}{\lambda} + 2 + N\mathcal{M}_\lambda(\rho_{t_k}) \right) e^{-2N(\varepsilon\alpha)^2} + 8NK e^{-\frac{\beta}{2} \min(\beta, \lambda)} \\ &\leq \left(\frac{t}{\Delta t} + 1\right) \left(\frac{4\kappa_t N(\varepsilon\alpha)}{\lambda} + 2 + 2Ne^{\lambda(\frac{1}{2} + \lambda)t} \mathcal{M}_\lambda^{x,v}(f_0) + 8N \right) e^{-2N(\varepsilon\alpha)^2} \\ &= \left(\frac{(2\kappa_t + \gamma)(c_\lambda + \beta)t}{\varepsilon\gamma} + 1 \right) \left(\frac{2\kappa_t(\varepsilon\gamma)}{\lambda} + 2e^{\lambda(\frac{1}{2} + \lambda)t} \mathcal{M}_\lambda^{x,v}(f_0) + 10 \right) N e^{-\frac{1}{2}N(\varepsilon\gamma)^2}. \end{aligned}$$

Using that by assumption $\varepsilon\gamma \leq \frac{1}{2}\kappa_t$, and the expression of β , leads to with

$$\mathbb{P}\left(\sup_{s \in [0, t]} \|\rho_s^N\|_{\infty, \varepsilon} \geq \kappa_t + \gamma\right) \leq \mathbf{C}_1 \left(1 + t \frac{2\kappa_t + \gamma}{\varepsilon\gamma} \left(c_\lambda + \sqrt{N}\varepsilon\gamma \max(1, \sqrt{N}\varepsilon\gamma\lambda^{-1})\right)\right) N e^{-\frac{1}{2}N(\varepsilon\gamma)^2},$$

$\mathbf{C}_1 := 10 + \kappa_t^2 \lambda^{-1} + 2e^{\lambda(\frac{1}{2} + \lambda)t} \mathcal{M}_\lambda^{x,v}(f_0)$. The conclusion follows using that $\lambda N^{-1/2} \leq \varepsilon\gamma$ by assumption.

6 Proof of Theorem 4

6.1 MKW estimates on deviation between particle and coupled systems

The first step of the proof is similar to the first step of the proof of Theorem 3. Precisely we start with equation (34), which reads now

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \sup_{s \in [0, t]} |X_{i,s}^N - Y_{i,s}^N| &\leq \int_0^t \frac{1}{N} \sum_{i=1}^N |V_{i,s}^N - W_{i,s}^N| ds \\ \frac{1}{N} \sum_{i=1}^N \sup_{s \in [0, t]} |V_{i,s}^N - W_{i,s}^N| &\leq \int_0^t \frac{1}{N^2} \sum_{i,j=1}^N |K(X_{i,s}^N - X_{j,s}^N) - K(Y_{i,s}^N - Y_{j,s}^N)| ds + \frac{t}{N-1} + \int_0^t \Lambda_s^N ds, \\ \text{where } \Lambda_s^N &:= \frac{1}{N} \sum_{i=1}^N \Lambda_{i,s}^N ds = \frac{1}{N} \sum_{i=1}^N \left| \frac{1}{N-1} \sum_{j=1}^N K(Y_{i,s}^N - Y_{j,s}^N) - \int_{\mathbb{R} \times \mathbb{R}} K(Y_{i,s}^N - x) \mu_s(dx, dv) \right|, \end{aligned}$$

since the initial conditions are now equal. Remark that if we introduce σ, τ two independent random variables with uniform law on $\{1, 2, \dots, N\}$, then the sum involving K becomes

$$\frac{1}{N^2} \sum_{i,j=1}^N |K(X_{i,s}^N - X_{j,s}^N) - K(Y_{i,s}^N - Y_{j,s}^N)| = \mathbb{E}_{\sigma, \tau} \left[|K(X_{\sigma,s}^N - X_{\tau,s}^N) - K(Y_{\sigma,s}^N - Y_{\tau,s}^N)| \right],$$

where we emphasize that the expectation is taken only with respect to (σ, τ) . So, we are in position to apply point (iii) Lemma 4 ; *i.e.* the part involving discrete uniform norms, and get:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \sup_{s \in [0, t]} |V_{i,s}^N - W_{i,s}^N| &\leq 8 \int_0^t \|\rho_s^N\|_{\infty, \varepsilon} \left(\frac{1}{N} \sum_{i=1}^N |X_{i,s}^N - Y_{i,s}^N| + \frac{\varepsilon}{2} \right) ds + \frac{t}{N-1} + \int_0^t \Lambda_s^N ds \\ &\leq \int_0^t (1 + 8\|\rho_s^N\|_{\infty, \varepsilon}) \left(\frac{1}{N} \sum_{i=1}^N |X_{i,s}^N - Y_{i,s}^N| + \frac{\varepsilon}{2} + \frac{1}{N-1} + \sup_{u \leq t} \Lambda_u^N \right) ds \end{aligned}$$

where ρ_s^N is the empirical measure of the $(Y_{i,s}^N)_{1 \leq i \leq N}$: $\rho_s^N := \frac{1}{N} \sum_{i=1}^N \delta_{Y_{i,s}^N}$. Applying finally Gronwall's Lemma on the interval $[0, t]$ where the quantity $\sup_{u \leq t} \Lambda_u^N$ may be considered as fixed, we get:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \sup_{s \in [0, t]} (|X_{i,s}^N - Y_{i,s}^N| + |V_{i,s}^N - W_{i,s}^N|) &\leq \left(\frac{\varepsilon}{2} + \frac{1}{N-1} + \sup_{s \leq t} \Lambda_s^N \right) \exp \left(t + 8 \int_0^t \|\rho_s^N\|_{\infty, \varepsilon} ds \right), \end{aligned} \quad (38)$$

$$\text{where } \Lambda_s^N := \frac{1}{N} \sum_{i=1}^N \left| \frac{1}{N-1} \sum_{j=1}^N K(Y_{i,s}^N - Y_{j,s}^N) - \int_{\mathbb{R} \times \mathbb{R}} K(Y_{i,s}^N - x) \mu_s(dx, dv) \right|. \quad (39)$$

We now focus on finding some concentration inequalities for the random variable $\sup_{t \in [0, T]} \Lambda_t^N$. In order to prove concentration inequalities for these supremum in time, we follow the same steps as in the proof of Proposition 2. Once it is be done, we will combine them with the concentration inequalities on $\sup_{t \in [0, T]} \|\rho_t^N\|_{\infty, \varepsilon}$ given by Proposition 2, and we will obtain some deviation upper bounds for

$$\frac{1}{N} \sum_{i=1}^N \sup_{s \in [0, t]} (|X_{i,s}^N - Y_{i,s}^N| + |V_{i,s}^N - W_{i,s}^N|).$$

6.2 Estimation of the fluctuations term Λ_t^N at fixed time t

We first establish the

Lemma 11. *Let $(Y_i^N)_{1 \leq i \leq N}$ be i.i.d. random variables, with a diffuse common law (i.e. a law that does not charge any atom). For $i = 1, \dots, N$ define the random variable Λ_i^N as follows:*

$$\Lambda_i^N = \frac{1}{N} \sum_{j=1}^N \left| \frac{1}{N-1} \sum_{j \neq i} K(Y_i^N - Y_j^N) - \mathbb{E}[K(Y_i^N - Y_j^N) | Y_i^N] \right|$$

Then for all $\alpha > 0$,

$$\mathbb{P}(|\Lambda^N| \geq \alpha) \leq 2N e^{-2\alpha^2(N-1)}.$$

Proof. Step 1. A calculation with Y_1^N frozen. To begin, we “freeze” $Y_1^N = a$ and define

$$\Lambda_1^N(a) := \frac{1}{N-1} \sum_{j \geq 2} K(a - Y_j^N) - \mathbb{E}[K(a - Y_j^N)].$$

By the definition (2) of K , and since $\mathbb{P}(Y_j^N = a) = 0$ by assumption the random variable ζ_j^a defined by:

$$\zeta_j^a = K(a - Y_j^N) + \frac{1}{2} = \begin{cases} 0 & \text{si } a < Y_j^N \\ 1/2 & \text{si } a = Y_j^N \\ 1 & \text{si } a > Y_j^N \end{cases}$$

is a Bernoulli variable with parameter $p_a := \mathbb{P}(Y_j^N > a)$. But since $p_a - \frac{1}{2} = \mathbb{E}[K(a - Y_j^N)]$, we have $\sum_{j \geq 2} \zeta_j^a - (N-1)p_a = \Lambda_1^N(a)$. So, by an application of Lemma 8 to the binomial variable $\sum_{j \leq 2} \zeta_j^a$:

$$\mathbb{P}(|\Lambda_1^N(a)| \geq \alpha) = \mathbb{P}\left(\left|\sum_{j \neq i} \zeta_j^a - (N-1)p_a\right| \geq (N-1)\alpha\right) \leq 2e^{-\alpha(N-1)\alpha^2},$$

Step 2. Summing up on N . Using the notation introduced in the previous step, we can rewrite $\Lambda^N = \frac{1}{N} \sum_{i=1}^N |\Lambda_i^N(Y_i^N)|$, and then

$$\begin{aligned} \mathbb{P}(\Lambda^N \geq \alpha) &\leq \mathbb{P}\left(\sup_{i=1, \dots, N} |\Lambda_i^N(Y_i^N)| \geq \alpha\right) \leq \sum_{i=1}^N \mathbb{P}(|\Lambda_i^N(Y_i^N)| \geq \alpha) \\ &= N \mathbb{E}\left[\mathbb{P}(|\Lambda_1^N(Y_1^N)| \geq \alpha | Y_1^N)\right], \end{aligned}$$

where we have used the fact that the variables $(Y_i^N)_{1 \leq i \leq N}$ are exchangeable. But by independence of Y_1^N and $(Y_i^N)_{i \geq 2}$, we obtain using the previous step that $\mathbb{P}(|\Lambda_1^N(Y_1^N)| \geq \alpha | Y_1^N) \leq 2e^{-2(N-1)\alpha^2}$ and the conclusion follows. \square

6.3 Estimation on the supremum in time of the fluctuations term

Lemma 12. *Assume that (Y_i^N, W_i^N) are N independent copies of a solution (Y, W) to (3) of initial law f_0 satisfying $\mathcal{M}_\lambda^{x,v}(f_0) < +\infty$ for some $\lambda > 0$, and Λ_t^N is defined by (39). Then, with $\rho_s := \mathcal{L}(Y_s)$ $\kappa_t := \sup_{0 \leq s \leq t} \|\rho_s\|_\infty$ we have provided that $\varepsilon \leq \min\left(\frac{1}{16}, \frac{\lambda}{2}, \lambda^{-2}\right)$.*

$$\mathbb{P}\left(\sup_{s \in [0, t]} |\Lambda_s^N| \geq C_t \varepsilon\right) \leq (\varepsilon + t) \left(\frac{5}{\varepsilon} + 4\kappa_t N \lambda^{-1} + \frac{N}{\varepsilon} e^{\lambda t (\frac{1}{2} + \lambda)} \mathbb{E}\left[e^{\lambda(|Y_0| + |W_0|)}\right] \right) e^{-2N\varepsilon^2},$$

with $C_t := 36 + 80\kappa_t + (1 + 3\kappa_t) \frac{8}{\lambda} \ln \mathbb{E}\left[e^{\lambda|W_0|}\right]$.

Proof. Step 1. Bounding the time fluctuations. Let $(Z_t)_{t \geq 0} = (Y_t, W_t)_{t \geq 0}$ be a solution to the non linear SDE (3). We show in that step that for $N \geq 2$:

$$\begin{aligned} \sup_{s \in [t, t + \Delta t]} |\Lambda_s^N - \Lambda_t^N| &\leq \left(16 \|\rho_t^N\|_{\infty, \varepsilon} + 4 \|\rho_t\|_{\infty} \right) \sup_{s \in [t, t + \Delta t]} \frac{1}{N} \sum_{i=1}^N |Y_{i,s}^N - Y_{i,t}^N| \\ &\quad + 4 \|\rho_t^N\|_{\infty, \varepsilon} |s - t| \left(\mathbb{E}[|W_0|] + 2 \right) + 10 \|\rho_t^N\|_{\infty, \varepsilon} \varepsilon. \end{aligned} \quad (40)$$

Indeed, by the definition (39) of Λ_s^N :

$$\begin{aligned} |\Lambda_N^s - \Lambda_N^t| &\leq \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1}^N |K(Y_{i,s}^N - Y_{j,s}^N) - K(Y_{i,t}^N - Y_{j,t}^N)| \\ &\quad + \frac{1}{N} \sum_{i=1}^N \left| \int_{\mathbb{R}} K(Y_{i,s}^N - x) \rho_s(dx) - \int_{\mathbb{R}} K(Y_{i,t}^N - x) \rho_t(dx) \right| \end{aligned}$$

Using the second point of Lemma 4 with two copies of a vector of joint law $\frac{1}{N} \sum_i \delta_{(Y_{i,t}^N, Y_{i,s}^N)}$, we may bound the first term in the r.h.s. by

$$\frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1}^N |K(Y_{i,s}^N - Y_{j,s}^N) - K(Y_{i,t}^N - Y_{j,t}^N)| \leq 8 \frac{N}{N-1} \|\rho_t^N\|_{\infty, \varepsilon} \left(\frac{1}{N} \sum_{i=1}^N |Y_{i,s}^N - Y_{i,t}^N| + \frac{\varepsilon}{2} \right).$$

To estimate the second term in the r.h.s, we use the third point of Lemma 4, applied to independent couples: The first one with law $\frac{1}{N} \sum_i \delta_{(Y_{i,t}^N, Y_{i,s}^N)}$ and (Y_t, Y_s) . It leads to the estimate

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \left| \int_{\mathbb{R}} K(Y_{i,s}^N - x) \rho_s(dx) - \int_{\mathbb{R}} K(Y_{i,t}^N - x) \rho_t(dx) \right| \\ \leq 4 \left(\|\rho_t\|_{\infty} \frac{1}{N} \sum_{i=1}^N |Y_{i,s}^N - Y_{i,t}^N| + \|\rho_t^N\|_{\infty, \varepsilon} \left(\mathbb{E}[|Y_s - Y_t|] + \varepsilon/2 \right) \right). \end{aligned}$$

Putting these two estimates together, using point (iv) of Proposition 3 in order to bound $\mathbb{E}[|Y_t - Y_s|]$ and taking the supremum in time leads to (40).

Step 2. Controlling the deviation. We define $t_k = k\varepsilon$ for $0 \leq k \leq k_M := \lfloor t\varepsilon^{-1} \rfloor$ and recall that by assumption $\varepsilon \leq \min\left(\frac{1}{16}, \frac{\lambda}{2}, \lambda^{-2}\right)$. Consider the events

$$\begin{aligned} \Omega_1 &:= \left\{ \|\rho_{t_k}^N\|_{\infty, \varepsilon} \leq \kappa_t + 1 \text{ for all } 0 \leq k \leq k_M \right\}, \\ \Omega_2 &:= \left\{ \sup_{s \in [t_k, t_k + \varepsilon]} \frac{1}{N} \sum_{i=1}^N |Y_{i,s}^N - Y_{i,t_k}^N| \leq \varepsilon(c_\lambda + 2\varepsilon) \text{ for all } 0 \leq k \leq k_M \right\}, \\ \Omega_3 &:= \left\{ |\Lambda_{t_k}^N| \leq 4\varepsilon \text{ for all } 0 \leq k \leq k_M \right\}. \end{aligned}$$

By Lemma 9 and the point (ii) of Lemma 3,

$$\mathbb{P}(\Omega_1^c) \leq (k_M + 1) \left(4\kappa_t N \varepsilon \lambda^{-1} + 2 + N e^{\lambda t \left(\frac{1}{2} + \lambda \right)} \mathbb{E} \left[e^{\lambda(|Y_0| + |W_0|)} \right] \right) e^{-2N\varepsilon^2}.$$

By Corollary 10, $\mathbb{P}(\Omega_2^c) \leq (k_M + 1) e^{-2N\varepsilon^2}$. By Lemma 11, $\mathbb{P}(\Omega_3^c) \leq 2(k_M + 1) e^{-8(N-1)\varepsilon^2} \leq 2(k_M + 1) e^{-2N\varepsilon^2}$ if $N \geq 2$. When the three above events are realized, we get with the help of the

bound (40)

$$\begin{aligned}
\sup_{s \in [0, t]} |\Lambda_s^N| &\leq \sup_{k=0, \dots, k_M} |\Lambda_{t_k}^N| + \sup_{k=0, \dots, k_M, s \in [t_k, t_k + \varepsilon]} |\Lambda_s^N - \Lambda_{t_k}^N| \\
&\leq 4\varepsilon + (20\kappa_t + 4)\varepsilon(c_\lambda + \varepsilon) + 4(\kappa_t + 1)\varepsilon \left(\mathbb{E}[|W_0|] + 2 \right) + 10(\kappa_t + 1)\varepsilon \\
&\leq \mathbf{C}'_t \varepsilon,
\end{aligned}$$

with $\mathbf{C}'_t := 26 + 4c_\lambda + 4\mathbb{E}[|W_0|] + \kappa_t(20c_\lambda + 4\mathbb{E}[|W_0|] + 30)$. Using the expression of c_λ and the fact that $\lambda \mathbb{E}[|W_0|] \leq \ln E[e^{\lambda|W_0|}]$, it is not difficult to show that $\mathbf{C}'_t \leq \mathbf{C}_t$, where \mathbf{C}_t is the constant introduced in Lemma 12. Moreover, gathering the three previous estimates on $\mathbb{P}(\Omega_i^c)$ for $i = 1, 2, 3$, we get

$$\mathbb{P}(\Omega_1^c \cup \Omega_2^c \cup \Omega_3^c) \leq \frac{\varepsilon + t}{\varepsilon} \left(5 + 4\kappa_t N \varepsilon \lambda^{-1} + N e^{\lambda t \left(\frac{1}{2} + \lambda \right)} \mathbb{E} \left[e^{\lambda(|Y_0| + |W_0|)} \right] \right) e^{-2N\varepsilon^2},$$

which concludes the proof. \square

6.4 Conclusion of the proof of Theorem 4

We now consider the two events

$$\Omega_1 := \left\{ \sup_{s \in [0, t]} |\Lambda_s^N| \geq \mathbf{C}_t \varepsilon \right\}, \quad \Omega_2 := \left\{ \sup_{s \in [0, t]} \|\rho_s^N\|_{\infty, \varepsilon} \geq \kappa_t + 1 \right\},$$

where \mathbf{C}_t is the constant defined in 12. By Lemma 12, and since we also assume $\varepsilon \geq \lambda N^{-1/2}$

$$\begin{aligned}
\mathbb{P}(\Omega_1^c) &\leq (\varepsilon + t) \left(\frac{5}{\varepsilon} + 4\kappa_t N \lambda^{-1} + \frac{N}{\varepsilon} e^{\lambda t \left(\frac{1}{2} + \lambda \right)} \mathbb{E} \left[e^{\lambda(|Y_0| + |W_0|)} \right] \right) e^{-2N\varepsilon^2}, \\
&\leq (\varepsilon + t) \lambda^{-1} N^{\frac{3}{2}} \left(5 + 4\kappa_t + e^{\lambda t \left(\frac{1}{2} + \lambda \right)} \mathbb{E} \left[e^{\lambda(|Y_0| + |W_0|)} \right] \right) e^{-2N\varepsilon^2}
\end{aligned}$$

and by Proposition 2, provided that $\varepsilon \leq 5\kappa_t \min\left(\frac{1}{16}, \frac{\lambda}{2}, \lambda^{-2}\right)$, and since we also assume $\varepsilon \geq \lambda N^{-1/2}$, the following bound holds with $\mathbf{D}_t := 10 + \kappa_t^2 \lambda^{-1} + e^{\lambda \left(\frac{1}{2} + \lambda \right) t} \mathbb{E} \left[e^{\lambda(|Y_0| + |W_0|)} \right]$

$$\begin{aligned}
\mathbb{P}(\Omega_2^c) &\leq \mathbf{D}_t \left(1 + t \frac{2\kappa_t + 1}{\varepsilon} \left(c_\lambda + \sqrt{N} \varepsilon \max(1, \sqrt{N} \varepsilon \lambda^{-1}) \right) \right) N e^{-\frac{1}{2} N \varepsilon^2}, \\
&\leq \mathbf{D}_t \left(1 + t \frac{2\kappa_t + 1}{\varepsilon} \left(c_\lambda + N \varepsilon^2 \lambda^{-1} \right) \right) N e^{-\frac{1}{2} N \varepsilon^2}, \\
&\leq \mathbf{D}_t \lambda^{-1} \left(\varepsilon + t(2\kappa_t + 1) \left(c_\lambda + \sqrt{N} \varepsilon \right) \right) N^{\frac{3}{2}} e^{-2N\varepsilon^2}.
\end{aligned}$$

If the two events are satisfied, then thanks to (38) we get

$$\begin{aligned}
\frac{1}{N} \sum_{i=1}^N \sup_{s \in [0, t]} (|X_{i,s}^N - Y_{i,s}^N| + |V_{i,s}^N - W_{i,t}^N|) &\leq \left(\left(\mathbf{C}_t + \frac{1}{2} \right) \varepsilon + \frac{1}{N-1} \right) \exp \left(t + 8 \int_0^t \|\rho_s\|_{\infty} ds \right), \\
&\leq \left(\mathbf{C}_t + \frac{1}{2} + \frac{2}{\lambda} \right) \exp \left(t + 8 \int_0^t \|\rho_s\|_{\infty} ds \right) \varepsilon,
\end{aligned}$$

where we have used that $\varepsilon \geq \lambda N^{-1/2} \geq \lambda(2(N-1))^{-1}$ when $N \geq 2$. which concludes the proof since

$$\mathbf{B}_t := \mathbf{C}_t + \frac{1}{2} + \frac{2}{\lambda}.$$

Defining also

$$\mathbf{A}_t'' := \lambda^{-1} \left(\mathbf{D}_t (1 + (2\kappa_t + 1)c_\lambda) + 5 + 4\kappa_t + e^{\lambda t(\frac{1}{2} + \lambda)} \mathbb{E} \left[e^{\lambda(|Y_0| + |W_0|)} \right] \right), \quad \mathbf{A}_t' := \lambda^{-1} \mathbf{D}_t (2\kappa_t + 1),$$

we obtain from the previous bound $\mathbb{P}(\Omega_1^c \cup \Omega_2^c) \leq (t + \varepsilon)(\mathbf{A}_t + \mathbf{A}_t' \sqrt{N\varepsilon}) N^{\frac{3}{2}} e^{-2N\varepsilon^2}$, and that concludes the proof. In particular it can be checked that $\mathbf{A}_t'' \leq \mathbf{A}_t$ where \mathbf{A}_t is defined by

$$\begin{aligned} \mathbf{A}_t &:= \lambda^{-1} \left[12 + \kappa_t^2 \lambda^{-1} + 2 e^{\lambda(\frac{1}{2} + \lambda)t} \mathbb{E} \left[e^{\lambda(|Y_0| + |W_0|)} \right] \right] (1 + (2\kappa_t + 1)c_\lambda), \\ \mathbf{A}_t' &:= \lambda^{-1} \left[10 + \kappa_t^2 \lambda^{-1} + e^{\lambda(\frac{1}{2} + \lambda)t} \mathbb{E} \left[e^{\lambda(|Y_0| + |W_0|)} \right] \right] (2\kappa_t + 1), \\ \mathbf{B}_t &:= 37 + \frac{2}{\lambda} + 80\kappa_t + (1 + 3\kappa_t) \frac{8}{\lambda} \ln \mathbb{E} \left[e^{\lambda|W_0|} \right]. \end{aligned}$$

7 Appendix

Proposition 4. *Let $K : (t, x) \in \mathbb{R}^+ \times \mathbb{R} \mapsto K_t(x) \in \mathbb{R}$ be a function belonging to $L_{loc}^\infty(\mathbb{R}^+, C^k(\mathbb{R}))$ for all $k \in \mathbb{N}$ and to $C^1(\mathbb{R}^+ \times \mathbb{R}^2)$. Consider the unique (in the class of measures) solution to the following linear PDE*

$$\partial_t f_t + v \partial_x f_t + K_t(x) \partial_v f_t = \partial_v (\partial_v f_t + v f_t), \quad (41)$$

for an initial condition $f_0 \in \mathcal{P}(\mathbb{R}^2)$ satisfying $\partial_x^k \partial_v^l f_0 \in L^2(\mathbb{R}^2)$ for all $k, l \in \mathbb{N}$. Then $f \in C^1(\mathbb{R}^+ \times \mathbb{R}^2)$ and is even two times continuously differentiable in (x, v) .

Proof. Differentiating equation (41) k times in x variable, and l times in v variable leads to

$$\begin{aligned} \partial_t (\partial_x^k \partial_v^l f_t) + v \partial_x (\partial_x^k \partial_v^l f_t) + K_t(x) \partial_v (\partial_x^k \partial_v^l f_t) - \partial_v (v \partial_x^k \partial_v^l f_t) - \partial_v^2 (\partial_x^k \partial_v^l f_t) \\ = - \sum_{k'=0}^{k-1} \binom{k}{k'} \partial_x^{k-k'} K_t(x) (\partial_x^{k'} \partial_v^{l+1} f_t) - \partial_x^{k+1} \partial_v^{l-1} f_t + l \partial_x^k \partial_v^l f_t, \end{aligned} \quad (42)$$

with the convention that a term containing a derivative with a negative power vanishes. Multiplying the above equation by $\partial_x^k \partial_v^l f_t$ and performing an integration by part, we obtain

$$\begin{aligned} \frac{d}{dt} \|\partial_x^k \partial_v^l f_t\|_2^2 &\leq \|\partial_x^k \partial_v^l f_t\|_2 \|\partial_x^{k+1} \partial_v^{l-1} f_t\|_2 + C_{t,k} \sum_{k'=0}^{k-1} \binom{k}{k'} \|\partial_x^{k'} \partial_v^{l+1} f_t\|_2 \|\partial_x^k \partial_v^l f_t\|_2 + \left(l + \frac{1}{2}\right) \|\partial_x^k \partial_v^l f_t\|_2^2, \\ \text{where } C_{t,k} &= \sup_{k'=1, \dots, k} \|\partial_x^{k'} K_t\|_\infty. \end{aligned}$$

Then summing these inequalities over the k, l such that $k + l \leq m$, we find

$$\frac{d}{dt} H_m(t) \leq \tilde{C}_{m,t} H_m(t), \quad \text{where } H_m(t) = \sum_{k+l \leq m} \left\| \partial_x^k \partial_v^l f_t \right\|_2^2,$$

and the constant $\tilde{C}_{m,t}$ is locally bounded in time. Since by assumption $H_m(0) < \infty$ for any $m \in \mathbb{N}$, we get that for any time $t \geq 0$, $\sup_{s \in [0, t]} H_m(s) < \infty$. Then by Morrey's inequality for m large enough ($m = 4$), $f \in L^\infty([0, t], C^2(\mathbb{R}^2))$, and so does $\partial_x^k \partial_v^l f$ for all $k, l \geq 0$. Using (42), we deduce that $\partial_t \partial_x^k \partial_v^l f \in L_{loc}^\infty(\mathbb{R}^+ \times \mathbb{R}^2)$, and in particular $\partial_x^k \partial_v^l f$ is continuous. So f is two times continuously differentiable in (x, v) . And using finally (41), we see that $\partial_t f$ is itself continuous in all the variables, as a sum of continuous functions. This concludes the proof. \square

References

- [1] R.J. Berman and Onnheim M. Propagation of chaos, Wasserstein gradient flows and toric Kähler-Einstein metrics. *arXiv:1501.07820*, 2015.
- [2] R. Catellier and M. Gubinelli. Averaging along irregular curves and regularisation of odes. *arXiv:1205.1735*, 2012.
- [3] E. Cépa and D. Lépingle. Diffusing particles with electrostatic repulsion. *Probab. Theory Related Fields*, 107(4):429–449, 1997.
- [4] M. Cullen, W. Gangbo, and G. Pisante. The semigeostrophic equations discretized in reference and dual variables. *Arch. Ration. Mech. Anal.*, 185(2):341–363, 2007.
- [5] A. M. Davie. Uniqueness of solutions of stochastic differential equations. *Int. Math. Res. Not.*, 2007(24):26, 2007.
- [6] W. Feller. An introduction to probability theory and its applications. Vol II. 2nd ed. Wiley Series in Probability and Mathematical Statistics. New York etc.: John Wiley and Sons, Inc. XXIV, 669 p. (1971)., 1971.
- [7] A. Figalli. Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.*, 254(1):109–153, 2008.
- [8] N. Fournier and A. Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields*, 162:707–738, 2015.
- [9] N. Fournier and M. Hauray. Propagation of chaos for the Landau equation with moderately soft potentials. *arXiv:1501.01802*, To appear in *Ann. Probab.*.
- [10] M. Hauray. Mean field limit for the one dimensional Vlasov-Poisson equation, Séminaire Laurent Schwarz, *arXiv:1309.2531*.
- [11] P.-E. Jabin and Z. Hauray. Mean Field Limit and Propagation of Chaos for Vlasov Systems with Bounded Forces, *arXiv:1511.03769*.
- [12] M. Hauray and S. Mischler. On Kac’s chaos and related problems. *J. Funct. Anal.* 266:6055–6157, 2014.
- [13] H. P. McKean, Jr. Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57.
- [14] A.-S. Sznitman. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, volume 1464, chapter Lecture Notes in Math., pages 165–251. Springer, Berlin, 1991.
- [15] M. Trocheris. On the derivation of the one-dimensional Vlasov equation. *Transport Theory Statist. Phys.*, 15:597–628, 1986.
- [16] A. Y. Veretennikov. On strong solutions and explicit formulas for solutions of stochastic integral equations. *Math. USSR, Sb.*, 39:387–403, 1981.